

6-1-2009

Preferred color correction for mixed taking-illuminant placement and cropping

Erin P. Fredericks

Follow this and additional works at: <http://scholarworks.rit.edu/theses>

Recommended Citation

Fredericks, Erin P., "Preferred color correction for mixed taking-illuminant placement and cropping" (2009). Thesis. Rochester Institute of Technology. Accessed from

This Thesis is brought to you for free and open access by the Thesis/Dissertation Collections at RIT Scholar Works. It has been accepted for inclusion in Theses by an authorized administrator of RIT Scholar Works. For more information, please contact ritscholarworks@rit.edu.

Preferred Color Correction for Mixed Taking-Illuminant Placement and Cropping

Erin P. K. Fredericks

B.S. Rochester Institute of Technology, NY (2007)

A thesis submitted in partial fulfillment of the requirements for the degree of Masters of Science in Imaging Science in the Chester F. Carlson Center for Imaging Science, College of Science, Rochester Institute of Technology

June 2009

Author _____

Accepted by _____

Coordinator, Graduate Degree Science Program

Date

CHESTER F. CARLSON CENTER FOR IMAGING SCIENCE
COLLEGE OF SCIENCE
ROCHESTER INSTITUTE OF TECHNOLOGY
ROCHESTER, NY

CERTIFICATE OF APPROVAL

M.S DEGREE THESIS

The Masters in Science Degree Thesis of Erin P. K. Fredericks has been examined and approved by the thesis committee as satisfactory for the thesis required for the Masters of Science Degree.

Dr. Mitchell R. Rosen,
Thesis Advisor

Dr. Mark D. Fairchild

Dr. James Ferwerda

Date

THESIS RELEASE PERMISSION FORM

CHESTER F. CARLSON CENTER FOR IMAGING SCIENCE
COLLEGE OF SCIENCE
ROCHESTER INSTITUTE OF TECHNOLOGY
ROCHESTER, NY

Preferred Color Correction for Mixed Taking-Illuminant Placement and Cropping

I, **Erin P. K. Fredericks**, hereby grant permission to Wallace Memorial Library of Rochester Institute of Technology to reproduce my thesis in whole or in part. Any reproduction will not be for commercial use or profit.

I additionally grant the Rochester Institute of Technology Digital Media Library (RIT DML) the non-exclusive license to archive and provide electronic access to my thesis in whole or in part in all forms of media in perpetuity. I retain all other ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Erin P. K. Fredericks

Date

Preferred Color Correction for Mixed Taking-Illuminant Placement and Cropping

Erin P. K. Fredericks

A thesis submitted in partial fulfillment of the requirements for the degree of Masters of Science in Imaging Science in the Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology

Abstract

The growth of automatic layout capabilities for publications such as photo books and image sharing websites enables consumers to create personalized presentations without much experience or the use of professional page design software. Automated color correction of images has been well studied over the years, but the methodology for determining how to correct images has almost exclusively considered images as independent indivisible objects. In modern documents, such as photo books or web sharing sites, images are automatically placed on pages in juxtaposition to others and some images are automatically cropped. Understanding how color correction preferences are impacted by complex arrangements has become important. A small number of photographs taken under a variety of illumination conditions were presented to observers both individually and in combinations. Cropped and uncropped versions of the shots were included. Users had opportunities to set preferred color balance and chroma for the images within the experiment. Analyses point toward trends indicating a preference for higher chroma for most cropped images in comparison to settings for the full spatial extent images. It is also shown that observers make different color balance choices when correcting an image in isolation versus when correcting the same image in the presence of a second shot taken under a different illuminant. Across 84 responses, approximately 60% showed the tendency to choose image white points that were further from the display white point when multiple images from different taking illuminants were simultaneously presented versus when the images were adjusted in isolation on the same display. Observers were also shown to preserve the relative white point bias of the original taking illuminants.

ACKNOWLEDGEMENTS

This work would not have been possible without the support of many.

To my faculty advisor, Dr. Mitchell Rosen and my committee, Dr. Mark D. Fairchild and Dr. James Ferwerda,

The faculty and staff of the Center for Imaging Science especially those in the Munsell Color Science Laboratory for their high degree of dedication and excellence,

To the MCSL support staff, Coleen Desimone and Valerie Hemlink, and the research staff David Wyble and Lawrence Taplin,

Generous support for my three research projects over the past two years from Thomas Lianza, Eastman Kodak Corporation, Microsoft with special thanks to Andrew Juenger for support and technical assistance, and the Chester F. Carlson family.

Frank Grum Memorial Scholarship for monetary support and inspiration in my studies,

To my fellow classmates Ben, Rodney, Stacey, Jonathan, Bingxin, Justin, Stefan, Shen, Mark, Nanette, and Marissa, for their support and guidance that helped me throughout the course of my studies by relating and sympathizing with my trials and tribulations,

Each of my observers whose patience and willingness created many insightful conclusions,

And to my family and friends, especially my parent and Jeremy, for their endless and unwavering support and encouragement without them I would not have realized this accomplishment.

Thank you all from the bottom of my heart.

June, 2009

Table of Contents

1. Introduction	1
2. Background	3
2.1 Color Reproduction	4
Hunt's objectives	4
Fairchild's Levels	6
Memory Colors and Color Reproduction	8
Simultaneous and Successive Color Matching	10
Cultural Difference	13
2.2 Color appearance	15
2.3 White balance	19
2.4 Summary of previous work	19
2.5 Camera Characterization	20
2.6 Display Characterization	24
3. Experimental Method	27
3.1 Purpose	27
Overview	27
Experimental I: Images in Isolation	28
Experimental II: Images Adjusted Simultaneously	28
3.2 Design	29
Method of Adjustment	29
Inspiration for Users Interface	29
3.3 Experimental Design	32
Experiment I: Images in Isolation	32
Experiment II: Multiple Images Presented Simultaneously	33
User Controls	35
Image Processing Workflow	37
User Instructions	40
3.4 Image Capture	40
Creating Apparent Distance	41
3.5 Display Specifications	42
3.6 Observer Population	43
4. Analysis & Discussion	44
4.1 System Characterization	44
4.2 Experiment I	48
Chroma Adjustment	48
White Point Adjustment	50
4.3 Experiment II	52
5. Conclusions	69
Future Works	74

List of Figures

FIG. 1-1. Examples of cohesive and disjointed images included in commercial brochures. Left document shows images with cohesiveness and the right shows disjointed images.....	2
FIG. 2-1. Schematic of the visual field.....	17
FIG. 2-2. Logarithm of luminance, or density, with the corresponding DC taken from a tone scale in three images each with different exposure setting. Black is 1 stop over exposed, magenta is the nominal setting and green is one stop under exposed.....	21
FIG. 2-3. Log luminance with the corresponding DC taken from a tone scale. Arrows represent the 0.3 logarithm of luminance unit shift.....	22
FIG. 2-4. Log luminance with the corresponding DC taken from a tone scale in three images each with different exposure settings that have been shifted so they line up.....	23
FIG. 2-5. Log luminance with the corresponding DC that have been shifted so they line up and then were fit with a curve that is a single channel's LUT.....	23
FIG. 2-6. Flowchart of characterization procedure used in Day et al. model [Day04].....	26
FIG. 3-1. Flowchart for inspiration experiment where the orange region is the processing that occurred prior to running the experiment and the blue region is the part that was preformed during the experiment	31
FIG. 3-2. Video experiments user interface.....	31
FIG. 3-3. Experiment I Part 3's user interface, which included both the white point and chroma user controls.....	33
FIG. 3-4. Experiment II Part 2 user interface, which includes the white point controls for both the left and right hand images	34
FIG. 3-5. Chromaticity plot that includes display primaries (circles) and white point (diamond) and the measured white points for each illuminant along with the Planckian locus function (gray curve) limited to the range (indicated by black dashed lines) used in the white point control.....	36
FIG. 3-6. Image processing workflow for this research.....	38
FIG. 3-7. Maximum luminance of the display for a given x chromaticity value (blue line) location of minimum is indicated by the red dashed lines.....	39
FIG. 3-8. Experiment images taken under three illumination conditions: (a) and (d) Daylight, (b) and (e) Fluorescent and (c) and (f) Tungsten. Images (a)-(c) are the long shots and (d)-(f) are the zoom shots.....	42

FIG. 4-1. Red, green and blue camera LUTs for each lighting condition. The line colors correspond to channels of the same color.....	44
FIG. 4-2. Comparison of neutral ramp (YW) to the sum of each primary ramp (YR + YG + YB) to evaluate additivity	46
FIG. 4-3. Chromaticity plot of primary ramps excluding DC lower than 25 for the display after characterization optimization.....	47
FIG. 4-4. Display LUTs between DC and radiometric scalars	47
FIG. 4-5. Relationship between the ΔC^* adjustment for the long and the zoom image for each taking illumination condition	50
FIG. 4-6. Results from the 28 observers for Experiment II Part 1 on the left and Part 2 on the right compared to the results in blue for Experiment I Part 1 when daylight is adjusted. The white star is the display white point and the blue star is the measured taking illuminant.....	54
FIG. 4-7. Cumulative probability percentage plot for the histogram in FIG. 4-6The white star corresponds to the display white point and the blue star to the scene measured daylight white point.	55
FIG. 4-8. Results from the 28 observers for Experiment II Part 1 on the left and Part 2 on the right compared to the results in blue for Experiment I Part 1 when adjusting fluorescent images. The white star is the display white point and the green star is the measured taking illuminant.....	56
FIG. 4-9. Cumulative probability percentage plot for the histogram in FIG. 4-8. The white star corresponds to the display white point and the green star to the scene measured fluorescent white point.....	57
FIG. 4-10. Results from the 28 observers for Experiment II Part 1 on the left and Part 2 on the right compared to the results in orange for Experiment I Part 1 where tungsten was adjusted. The white star is the display white point and the orange star is the measured taking illuminant.	58
FIG. 4-11. Cumulative probability percentage plot for the histogram in FIG. 4-10. The white star corresponds to the display white point and the orange star to the scene measured tungsten white point.	
FIG. 5-1. Average results for daylight-captured image for the white point adjustment.....	70
FIG. 5-2. Average results for fluorescent-captured image for the white point adjustment.....	71

FIG. 5-3. Average results for tungsten-captured image for the white point adjustment.....	72
FIG. 5-4. Depiction of difference between image adjusted in isolation, the blue cloud, and images adjusted simultaneously (on the same page), the red cloud. The magenta and cyan stars are the original white points for Image 1 and Image 2.....	73

List of Tables

Table 2-I. Hunt's color reproduction objectives	
Table 2-II. Fairchild's color reproduction levels	7
Table 3-I. Experimental design for Experiment I summary include the parts and the corresponding controls provided to the observer	32
Table 3-II. Experimental design for Experiment II summary include the parts and the corresponding controls provided to the observer	34
Table 4-I. Statistics for ColorChecker® verification of camera characterization under the three illumination conditions in terms of CIEDE2000	45
Table 4-II. Statistics for verification dataset for display characterization in CIEDE2000	48
Table 4-III. Apparent distance percentage results for chroma for each illuminant	50
Table 4-IV. CIEDE2000 between the display white point and the illuminant white point for both the scene measured and average adjusted white point.....	51
Table 4-V. P-value statistics for Experiment I when the daylight image white point was adjusted in isolation compared to both fluorescent and tungsten images adjusted in isolation	2
Table 4-VI. Standard deviation of x chromaticity data for the 28 observers when daylight was being adjusted.....	54
Table 4-VII. Standard deviation of x chromaticity data for the 28 observers when fluorescent was being adjusted.....	56
Table 4-VIII. Standard deviation of x chromaticity data for the 28 observers when tungsten was being adjusted.....	60
Table 4-IX. Ordered categorical results for the daylight image looking at the CIEDE2000 distance of the adjusted white point from the display white point. S (standalone) is from Experiment 1, Part 1 where the image was adjusted in isolation. F (fluorescent) and A (tungsten) are from Experiment 2 where the image was adjusted in the presence of the associated second image. Part 1 only allowed the first image to be adjusted. Part 2 allowed both image white points to be adjusted. Leftmost category is highest adjusted color difference from display white point, rightmost category is lowest adjusted color difference from display white point and bold line boxes represent multiple categories adjusted to preferred values that differ to within one CIEDE2000 unit.....	61
Table 4-X. Ordered categorical results for the fluorescent image looking at the CIEDE2000 distance of the adjusted white	

point from the display white point. S (standalone) is from Experiment 1, Part 1 where the image was adjusted in isolation. D (daylight) and A (tungsten) are from Experiment 2 where the image was adjusted in the presence of the associated second image. Part 1 only allowed the first image to be adjusted. Part 2 allowed both image white points to be adjusted. Leftmost category is highest adjusted color difference from display white point, rightmost category is lowest adjusted color difference from display white point and bold line boxes represent multiple categories adjusted to preferred values that differ to within one CIEDE2000 unit.....63

Table 4-XI. Ordered categorical results for the tungsten image looking at the CIEDE2000 distance of the adjusted white point from the display white point. S (standalone) is from Experiment 1, Part 1 where the image was adjusted in isolation. F (fluorescent) and D (daylight) are from Experiment 2 where the image was adjusted in the presence of the associated second image. Part 1 only allowed the first image to be adjusted. Part 2 allowed both image white points to be adjusted. Leftmost category is highest adjusted color difference from display white point, rightmost category is lowest adjusted color difference from display white point and bold line boxes represent multiple categories adjusted to preferred values that differ to within one CIEDE2000 unit..... 64

Table 4-XII. Combined categorical results taken from Table IX, Table X and Table XI according to the smallest CIEDE2000 adjustment difference from display white where gray boxes are all possible symbols.....65

Table XIII. Combined categorical results taken from Table IX, Table X and Table XI according to the largest CIEDE2000 adjustment difference from display white where gray boxes are all possible symbols.....66

Table 4-XIV. Percentage of observer responses maintaining relative order of adjusted white points for Isolation presentation and Simultaneous presentation and percentage of observer responses increasing the average white point distance from the display white point when comparing Isolation presentation to Simultaneous presentation.....67

Table 4-XV. Experiment II Part 1 p-values for each illuminant.....68

1. Introduction

On a path toward valuable improvements in image appearance models, an investigation was launched to better understand how image cropping affects image appearance and how multiple images in close proximity affect each other's appearance. This research is a first step toward creating image-processing algorithms for automated layout programs. Such programs would optimize the appearance of images after automatic cropping as well as, optimize the appearance of images placed in close proximity on a page but taken under different lighting conditions.

More specifically, the first question addressed in this investigation was, How will chroma preference change when images were cropped? The second question was, How will the preferred white point adjustment for images change when viewed in isolation versus viewing multiple images in juxtaposition? Answering these questions could lead to automated image editing system for various applications. Some examples of applications where collections of images are viewed as a cohesive piece are brochures, annual reports, multimedia content and the newly popular photo books. High quality examples of such output generally involve professionals manually manipulating images or in some cases are simply limited to stock images that are pre-manipulated to look the same. The first case may end up with an appealing result, but will be costly in both manpower and time. The latter case is shown by this study to be less than optimal

due to our finding that images adjusted in the presence of other images will almost always have a different preferred setting relative to their adjustment in isolation.

Figure 1-1 shows two examples of multiple image documents. In the one on the left, the images look to have been taken at different times under different conditions and yet they all appear to belong together. For the right example, the two images look disjointed.



FIG. 1-1. Examples of cohesive and disjointed images included in commercial brochures. Left document shows images with cohesiveness and the right shows disjointed images.

Today's growth of digital printing options enables consumers to create personalized documents without much experience or the use of professional layout software. Lulu is an example of a company profiting from this technology and business model [Phillips08]. Part of their success is attributed to the quality of their products leading to end consumer satisfaction. In this type of application, the quality of the output is a function of the images found on the pages of the submitted document. Adding image processing within the automated workflow to improve color in automatically cropped images and cohesiveness among multiple images presented on individual pages could improve the quality of the result.

Color management techniques that are already included in many printing workflows are typically blind to the issues of color preferences for cropping and simultaneous presentation of multiple images. The addition of these new techniques will further enhance the final product.

There are two topics that this research addresses; white point adjustment and chroma adjustment. How do multiple images taken under different illumination conditions and presented simultaneously affect the preferred white point adjustment of each of the images? How does cropping affect preferred chroma adjustment? Cropping is related to *apparent distance* because it can create the illusion of a shorter distance between the subject of an image and the camera that captured the image. This is expected to have an increase the preferred chroma of the image.

2. Background

2.1 Color Reproduction

Hunt's objectives

Color reproduction is often defined as the production of an image on a media different from the original that includes color. Reproduction methods are dependent on the intent of the relationship between the original and the reproduction. Hunt's [Hunt57] six objectives of color reproduction are: spectral, colorimetric, exact, equivalent, corresponding color and preferred. In Table 2-I a derived mathematical representation for each of his objectives is presented with the exception of his preferred objective.

Table 2-I. Hunt's color reproduction objectives

Objective	Derived mathematical representation
1. Spectral	$R_{\lambda,1} = R_{\lambda,2}$
2. Colorimetric	$XYZ_1 = XYZ_2 \quad L_1 \neq L_2$
3. Exact	$XYZ_1 = XYZ_2 \quad L_1 = L_2$
4. Equivalent	$JCh_1 = JCh_2$
5. Corresponding Color	$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_1 = \mathbf{M}_{CAT} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_2$ when $L_1 = L_2$
6. Preferred	No simple mathematical representations for this objective

Spectral reproduction is the precise match of the reflectance of the original object. This is optimal since viewing conditions can vary and the match will be maintained. When the same pigments are used in a controlled process, a reflectance match is possible. The problem is that in most applications where

pigments differ from the original, reproduction will result in a non-spectral match although researchers such as Urban [Urban07] and Rosen [Rosen06] are working to solve this issue. When a sufficient spectral match is not achieved, one of the other five objectives is appropriate for color reproduction.

The colorimetric reproduction objective is to match the relative tristimulus values of two samples. This means that the spectra for the original and the reproduction can be distinct and are matched via filtering through the human visual system. This type of match is known as a metameric pair. A change in the viewing conditions can cause the pair to no longer match. Often colorimetric reproduction will not match for different observers, thus, exhibiting observer metamerism. The large range of matches from a collection of observers is attributed to the variation in color matching functions within the human population as a result of many things including age. Colorimetric matches are commonly created when color-forming materials differ, such as those found in photography between a natural scene and an inkjet print.

The objective of exact reproduction is matching the absolute tristimulus values meaning the absolute luminance levels must also match. Exact reproduction results in a match in the adaptation state between the original and the reproduction. Tristimulus matching objectives assume that the illumination for the original and the reproduction are the same.

Equivalent reproduction objective is used when the color of the illumination of the original and the reproduction differ. A simple tristimulus match would be incorrect since appearance would not be maintained. Matching

using lightness, J , chroma, C , and hue angle, h , and a chromatic adaptation model will take the illumination condition into account. This objective is incorrect when the change in the color of the illumination is large. This is due to the inconsistencies in lightness and chroma as well as brightness and colorfulness. A change from blue illumination to orange is an example of a situation where equivalent reproduction would not be appropriate.

Corresponding color reproduction objective overcomes the shortcoming of equivalent reproduction by adjusting tristimulus values to create a color appearance match when the original and reproduction are under the same illumination levels.

Preferred color reproduction cannot be mathematically represented because it is defined where the colors depart from equality of appearance to those of the original in order to give a more pleasing image. These deviations are dependent on viewers' preference variability. The most appropriate method for consumer image applications according to Hunt is this color reproduction objective. Preferred color reproduction will be the focus of this paper.

Fairchild's Levels

Fairchild [Fairchild05] described levels of reproduction based on Hunt's objectives for modern color imaging systems. Summaries of the levels are included in Table 2-II. Each level must satisfy the requirements of the previous level to define a reproduction at the succeeding level.

Table 2-II. Fairchild's color reproduction levels.

Levels	Summary of definitions
1. Color	Simply reproduced color
2. Pleasing	Acceptable images
3. Colorimetric	Characterization of devices to produce known XYZ values and is only appropriate when the viewing conditions for the original and the reproduction are the same.
4. Color Appearance	Must have knowledge of the viewing conditions for both the original and the reproduction which are input into a color appearance transform along within known XYZ values for the original scene.
5. Color Preferred	Purposefully manipulating the color in a reproduction such that the result is preferable to the user over an accurate reproduction.

Level one, color reproduction is as simple as the name suggests, a reproduction that includes color. This level is a requirement for the rest of the levels, but is rarely an acceptable place to stop. Pleasing color reproduction, or level two, best describes the current state of consumer imaging. Pleasing reproduction makes a color reproduction acceptable and is achievable with trial and error within a system. Fedorovskaya et al. [Fedorovskaya97] found parabolic relationship between pleasing appearance and chroma adjustment. Slightly more chromatic was favored, but at very high levels of adjustment the naturalness decreases. Some applications require more accurate reproduction of the colors such as in color science experiments where known colorimetry is required. A colorimetric reproduction requires characterization of each component of the system enabling conversion between device digital values and tristimulus values. This method ensures that the color reproduced has known colorimetric values. Low system variability or appropriate calibration is required to maintain validity of the characterization. Color appearance reproduction, level four, uses color appearance models, viewing condition information for the original and the

reproduction and good calibration and accurate characterization of the entire system to produce a reproduction.

Preferred color reproduction, level five, produces aesthetically-preferred results, but it has not been achieved in consumer imaging due to the inability for consumer products to create colorimetric reproductions according to Fairchild [Fairchild05].

Consumer imaging system workflows do not generally account for the changes in viewing conditions that would be required to reach a color appearance reproduction. Uncharacterized devices in the consumer workflow prevent the reproduction from meeting the colorimetric color reproduction, level three, criteria. When digital images are processed and displayed or printed by consumers, they are often adjusted to produce a result that looks pleasing to the consumer, with less concern for how accurately the colors in the original scene are produced.

Memory Colors and Color Reproduction

Reproduction of skin-tones is a metric of subjective quality for the consumer digital imaging industry. Preferred reproductions of complexions are difficult to achieve due to basic physiological mechanisms, mainly memory color and chromatic adaptation.

Early work related to memory colors by Katz [Katz35] described the relationship between memory colors and common illumination. He introduced the term *common illumination* defined as daylight to explain the illumination condition of memory color perceptions. Under common illumination there is

strong agreement between actual color impression and memory color. Atypical illumination perceptions are dominated by the memory of the object's color appearance. He found that large deviations from daylight lead to a reduction in the effect of memory color on the perception of the object.

A later experiment by Newhall and Pugh [cf. Newhall57] concluded that when people indicated colors of familiar objects from memory, they reported colors with higher value and chroma than the objects' actual colors. Newhall, Burnham and Clark [Newhall57] attempted to test plausible causes for the increase in value and chroma for memory of colors by checking results against theories associated with adaptation and retention view. Incorporation of adaptation time with the luminance data tested luminance adaptation and was found insignificant. In two subsequent experiments, steps were taken to reduce observer adaptation, yet increases in value and chroma were still reported. The results of these three experiments did not support a theory that adaptation was the cause of these memory color phenomena. Retention of view theory holds that the dominant aspects of stimulation are remembered over the weaker characteristics and therefore exaggerated. *Chromatic impressiveness* was confirmed via experimentation where highly chromatic colors required nearly double the chroma to match and less impressive colors required significantly less. Retention of view was concluded as a most probable explanation for the findings of Newhall and co-workers' research.

Simultaneous and Successive Color Matching

Newhall's [Newhall57] theory has some important practical implications. His results showed that simultaneous and successive matching experiments differ in resulting remembered colors. In applications where two stimuli are presented successively, the application of simultaneous matching colorimetry was found to be inappropriate. Similarly, it was found for successive matching experiments simultaneous matching colorimetry should not be applied. In the photographic industry, the reproduction is seldom compared to the original scene due to its lack of availability. Thus, photo quality is usually judged based on memory. Within the context of images the objects themselves can provide information that contributes to the memory of the color. Color memory experiments are used in successive match experiments where the observer has to remember a stimulus and then generate a match. Colors that are frequently a part of visual experience are known as memory colors. Skin tones, blue sky and green grass are examples of memory colors. Results for color memory experiments can differ from those of memory color due to the added contextual information given by the object and frequency of experience by the observer [Newhall57, Bartleson61]. Further research by Siple and Springer [Siple83] showed that memory of a color and preferred hue are independent of contextual cues, such as shading or texture. The introduction of contextual information included in images will not affect the hue, but will impact the value and chroma.

Bodrogi and Tarczali [Bodrogi01] showed that memory color shifts were applicable in practical applications such as images. de Ridder [de Ridder96]

concluded that preferred reproductions of complex images were more chromatic than the original scene. This conclusion hinted at the need for higher color contrast, which was hypothesized to improve object recognition. Test images included a close-up portrait and a distant shot within a scene. The portrait image showed a linear relationship between colorfulness and saturation indicating a dependency on image lightness. The other image demonstrated an independent relationship with slight compression, which supports the use of saturation instead of chroma in images.

Bartleson [Bartleson59] bridged the gap between memory color and color reproduction when he found that chromaticities of preferred representations of flesh differed from those of natural. These results were also supported the work by Buck and Froelich [Buck48]. Bartleson found that reproduced chromaticities of Caucasian skin-tones in oil paintings, pastels, Carbro paints, Kodak Flexichrome and Dye Transfer prints were more chromatic than the originals. Chromaticity differences led to further research on understanding memory color and its effect on skin-tone reproductions. He also concluded that saturation and brightness are changed significantly in memory, and that hues were the same as those of the original. Bartleson [Bartleson59] recognized early on that his results applied to simple image configuration and but memory color changes with complexity, skin-tone area, variation in viewing and adaptation.

Pinney and DeMarsh [Pinney63] looked at preferred reproduction as it relates to observers' adaptation by investigating the *constant-appearance* and *gray-point concepts*. The constant-appearance concept is that preferred color

remains relatively constant whatever the adaptation. The experimental results confirmed that the concept of constant preferred color is maintained regardless of adaptation. This is similar to Katz's idea of a common illuminant where memory of a color does not change with the adaptation illumination. The gray-point concept that observers adapt to changes in illumination, therefore their concept of gray must also be changing. Experiments where observers adjusted the white point of slide taken under different illumination conditions resulted in warmer illumination slides had warmer gray-points and cooler slides had cooler gray-points.

Hunt, Pitt and Winter [Hunt74] attributed the previous finding of preferred reproductions of Caucasian skin to have been process-limited, referring to Bartleson's work on artwork. The "natural mean" or average measurements of skin color were not attainable due to chromaticity being outside the achievable range of the systems. Their research focused on looking at reflectance prints and projected transparencies. They found that green grass and Caucasian skin preference ellipses contained the "natural mean", but blue sky had a higher purity because the chromaticity of the mean was lower than the observers preferred chromaticities. Overall the preferred skin reproductions were more chromatic making it more yellow in both processes. Imai et al. [Imai98] showed that preferred reproduction for a patch differed from those of facial pattern. Kuang et al. [Kuang05] later found that illumination and scene content had a significant influence on the color reproduction of skin and grass in images. This supports the

idea that contextual cues and physiological processes are a factor in image reproduction.

Cultural Differences

Fernandez and Fairchild [Fernandez02] [Fernandez05] performed two experiments to better understand observer preferences and cultural differences in preferred color reproduction. Analysis looked at each experiment separately and then together. For the first experiment observers were asked to rank order sets of printed images from best to worst. Using Thurston's Law of Comparative Judgments the rank order results were converted to scales of preference. Cultural differences were found statistically significant, however implementation of these differences is impractical in most applications. Yano and Hashimoto [Yano97] work supported these conclusions since Asian skin tones differed from Caucasian skin tones. For the second experiment observers were asked to adjust the images to a preferred reproduction. This experiment was broken down into two phases. The first allowed the observers to manipulate any dimension in any order until they reach a preferred reproduction. The second phase was to manipulate a single dimension at a time to a preferred level. Variability within observer was half that of the variability between observers. It was found that the distance of the starting point from the end point would not provide insight into the preferred color reproduction, because the manipulation will be used to correct improper characterization. Images containing people were found to have the smallest deviation from the mean image. Crossover analysis was performed, which confirmed that there is little difference between the methods used in this research.

It was determined that global manipulation tool for hue rotation was difficult to judge and therefore did not have clear results.

As previously explained, preferred reproductions require first creating a color appearance match that includes compensation for the viewing conditions within the image workflows. Fairchild and Braun conducted early work comparing manual adjustments of images to color appearance models such as von Kries, Hunt model, CIELAB and RLAB [Fairchild97]. This work proved that psychophysical techniques of method of adjustment and pair-comparison were appropriate for use when improving color-appearance models and proving those improvements respectively. There were some important notes from this work that should be considered in future work. Observers should be instructed to shift their focus around the image in order to prevent local adaptation. The results of such experiments will have larger color differences and higher variability caused by image complexity and observer metamerism. Our research will use similar techniques and make similar considerations in order to further improvements to color appearance models.

A second motivation for this research is to collect information useful for automatic cropping techniques. Artists have been painting grand landscapes for hundreds of years. It is known that illusion of depth is enhanced with the use of color. Stelmack [Stelmack65] explored the effect of hue and chroma on apparent distance. He found that with small solid patches there was a significant effect on apparent distance with changes in hue and chroma where brighter hues with more chroma appeared closer. Unlike Stelmack, our research looked at images. We also

wish to understand how apparent distance impacts color appearance. Later Fairchild and Johnson [Fairchild99] found that the size of the region is a factor in the variance of preferred chroma. There are a number of algorithms out there that enhance the chroma of specific regions in a scene [Braun07][You08].

2.2 Color appearance

Chromatic adaptation uses cognitive and physiological mechanisms of the human visual system to map object colors to alternative white points [Fairchild05 (book)]. Cognitive mechanisms take into account knowledge of scene content. Physiological mechanisms take into account a change in sensitivities of the photoreceptors and the neurons first few stages in the visual pathway. Chromatic adaptation occurs when prolonged exposure to a colored stimulus reduces awareness of that color while viewing an environment. Imaging systems do not innately have either group of mechanisms. Thus, they require image processing to produce reproductions that have apparent matches to the original scene. White balancing algorithms within digital cameras apply image processing that somewhat mimics the physiological mechanisms.

Color appearance models depend on the stimulus and the nearby space or time. The spatial configuration of a stimulus is important because eyes are in constant motion. Temporal effects are generally not pertinent in common applications. There are four components of the visual field that are required to be defined in most color appearance models: stimulus, proximal field, background and surround. FIG. 2-1 is a schematic of the visual field [Fairchild05]. The

stimulus is a colored element that a color appearance measure is desired. Typically the stimulus is a uniform patch covering two-degrees of angular subtense. A uniform patch eliminates the problems associated with an inhomogeneous retina such as assumptions required to implement a standard color observer. When a stimulus is an image however some the assumptions can be invalid. Another factors is that when viewing real scenes an entire object as a uniform stimulus, which can be larger than a two-degree angular subtense.

The proximal field is the region in immediate vicinity of stimulus that extends about two-degrees from the edge of the stimulus in an image this would be the surrounding pixels. Generally the proximal field is equivalent to the background. The background extends from the edge of the stimulus about 10-degrees. The definition of the background is required in color appearance models to account for simultaneous contrast effects. For an image the background is usually made up of the area surrounding the image. The surround is the field outside the background generally considered the rest of the room where the stimulus is being viewed.

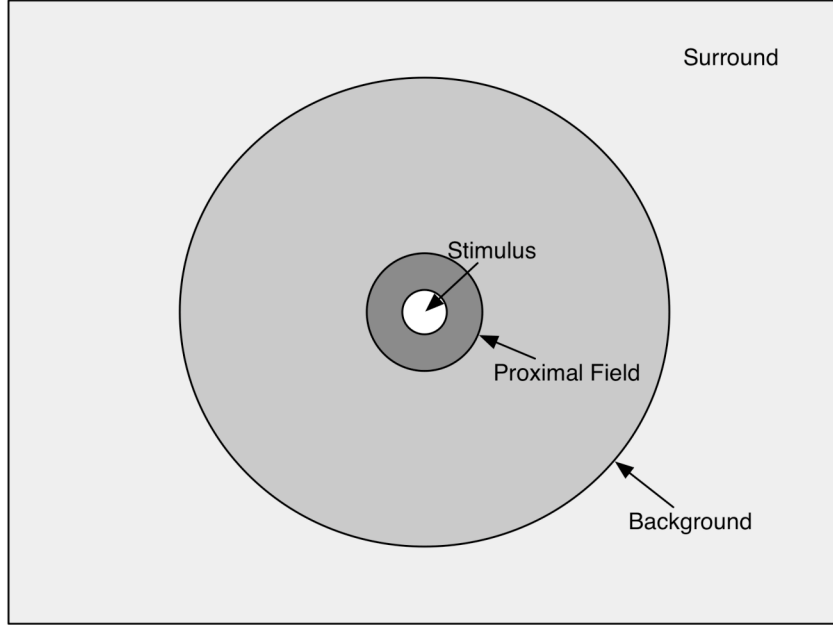


FIG. 2-1. Schematic of the visual field.

The simplest chromatic adaptation model is the von Kries model and is the basis for all other models. von Kries [von Kries02] considered his ideas to be an extension of Grassman's laws of color mixing under different viewing conditions. The model uses a simple gain-control, k_L , k_M and k_S , on the cone responses, L , M and S , shown in Eq. 2.1.

$$\begin{aligned}
 L_a &= k_L L \\
 M_a &= k_M M \\
 S_a &= k_S S
 \end{aligned}
 \tag{2.1}$$

The coefficients are calculated using Eq. 2.2, where L_{white} , M_{white} and S_{white} are the cone responses for the scene white.

$$\begin{aligned}
 k_L &= \frac{1}{L_{white}} \\
 k_M &= \frac{1}{M_{white}}
 \end{aligned}
 \tag{2.2}$$

$$k_s = \frac{1}{S_{white}}$$

Corresponding colors are calculated using Eq 2.3, where the ones are for the initial illumination condition and twos are for the second and final illumination condition.

$$\begin{aligned} L_2 &= \left(\frac{L_1}{L_{white,1}} \right) L_{white,2} \\ M_2 &= \left(\frac{M_1}{M_{white,1}} \right) M_{white,2} \\ S_2 &= \left(\frac{S_1}{S_{white,1}} \right) S_{white,2} \end{aligned} \quad (2.3)$$

The matrix representation of these equations is shown in Eq. 2.4, where \mathbf{M} is the transformation from CIE tristimulus values to relative cone responses.

$$\begin{bmatrix} X_2 \\ Y_2 \\ Z_2 \end{bmatrix} = \mathbf{M}^{-1} \begin{bmatrix} L_{white,2} & 0 & 0 \\ 0 & M_{white,2} & 0 \\ 0 & 0 & S_{white,2} \end{bmatrix} \begin{bmatrix} 1/L_{white,1} & 0 & 0 \\ 0 & 1/M_{white,1} & 0 \\ 0 & 0 & 1/S_{white,1} \end{bmatrix} \mathbf{M} \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \end{bmatrix} \quad (2.4)$$

The von Kries model is not only the simplest but it also is the basis of most other models. Proper estimation of the white point of the taking illumination is crucial to the acceptability of the final output. A simple von Kries-like transformation is typically used to convert from the estimated colorimetry to the standard condition in consumer imaging technology [Sharma03]. This method is computationally inexpensive and predicts a majority of chromatic adaptation phenomena. But, it does have some limitations. Jameson and Hurvich [Jameson61] found that this method is inaccurate and inadequate for predicting the appearance of color patches. Developments of better models have taken place, but these models tend

to be computationally cumbersome and require knowledge of the scene that cannot be collected via a camera. It is important to note that these corrections are based on image processing research done on images in isolation.

2.3 White balance

White balancing or color balancing is a color correction applied to an image to ensure that the neutrals do not have a colorcast. Without this color correction the resulting images from camera systems would have an undesirable cast that the photographer did not experience when taking the picture. The difference is a consequence of the camera system not adapting and discounting the illumination conditions the way the human visual system would. Prior to digital cameras, film sensitivities were designed for specific illuminations. Changing the sensitivities based on the illumination results in neutrals reproduced as neutrals. Consumer digital cameras use a chromatic adaptation transform, most commonly the von Kries, to account for adaptation [Sharma03]. Using a color appearance model requires the proper estimation of the chromaticities of the scene illumination. Jiang [Jiang03] summarized a variety of methods of estimating the illumination, which is important because the acceptability of a white balancing algorithm is highly dependent on the ability of the system to estimate the illumination.

2.4 Summary of previous work

Preferred color reproductions were found to differ from their originals. Memory colors such as skin tones, sky and grass have been found to be very influential in preferred reproductions. All the work in this area has looked at images in isolation. Preference experiments have been shown to have high variability from observer to observer, but there do not seem to be a significant difference between cultures. Not much work has looked at a chroma dependence on size or distance. Early work does look at the extremes, very far and very close, but does not address the middle distances.

This work begins the process of looking at these new issues. One question is whether preferred white points differ for an image when in the presence of other images taken under different illumination conditions compared to choices made when looking at the images in isolation. A second question is whether or not preferred chroma changes when the apparent distance of objects within the scene are changed.

2.5 Camera Characterization

A camera system can be characterized using a generic a look up tables (LUT) and matrix approach. The goal of the model is to approximately relate digital count values (DC) to the colorimetric quantities that the camera is capturing. The model consists of three one-dimensional LUT and an optimized 3x3 matrix. It is important to ensure that all possible DC are well handled by the

LUT. One method for probing how a camera handles high and low light levels is by changing the exposure settings for the camera as grayscale targets are photographed. An alternative method involves changing for each photograph the intensity of the light source illuminating the target. For the change in exposure method, measured tristimulus values are associated to a single (nominal) exposure and for other exposures, calculation of the luminance levels that would correspond to the DC under nominal conditions are calculated. DC from a single channel collected from a gray tone scale for three exposure sets are plotted against the logarithm of measured luminance, or density, from the nominal exposure level as seen in FIG. 2-2. The black dots are from a one stop over exposed image, the magenta dots are from the nominal exposure and the green are from one stop under exposed. The over and under exposed values are shifted horizontally in log luminance to overlap the nominal exposure, FIG. 2-3 illustrates the shifts.

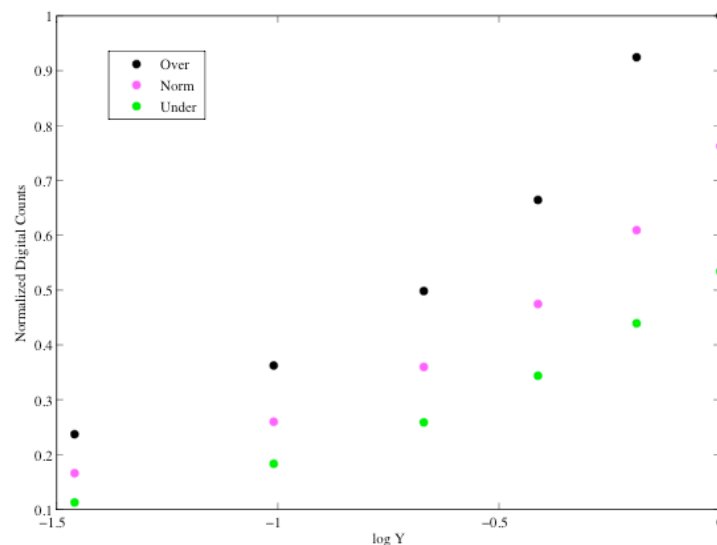


FIG. 2-2. Logarithm of luminance, or density, with the corresponding DC taken from a tone scale in three images each with different exposure setting. Black is 1 stop over exposed, magenta is the nominal setting and green is one stop under exposed.

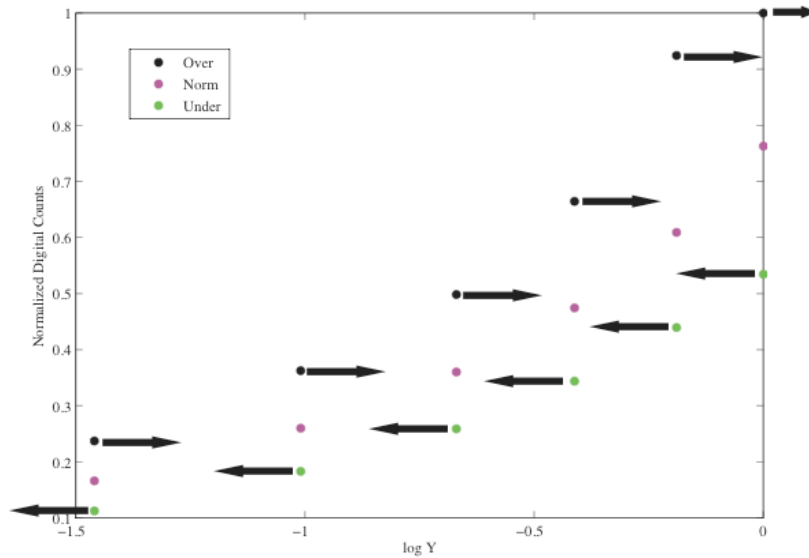


FIG. 2-3. Log luminance with the corresponding DC taken from a tone scale. Arrows represent the 0.3 logarithm of luminance unit shift.

One stop of exposure is approximately 0.3 log luminance units. The shifted points are shown in FIG. 2-4 moved this distance. While most of the dots overlap in the nominal region, some of the over exposed dots extend the curve into higher log luminances and some of the under exposed dots extend the curve into lower log luminances. A curve as shown in FIG. 2-5 as a blue curve is fit to the ensemble shifted values.

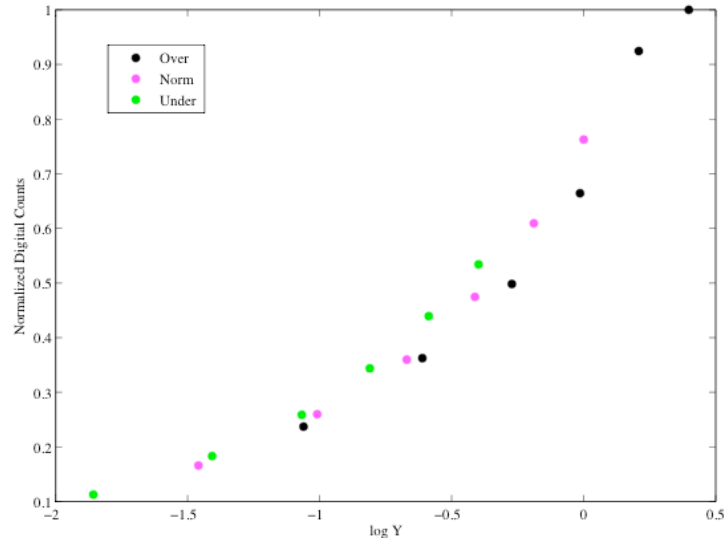


FIG. 2-4. Log luminance with the corresponding DC taken from a tone scale in three images each with different exposure settings that have been shifted so they line up.

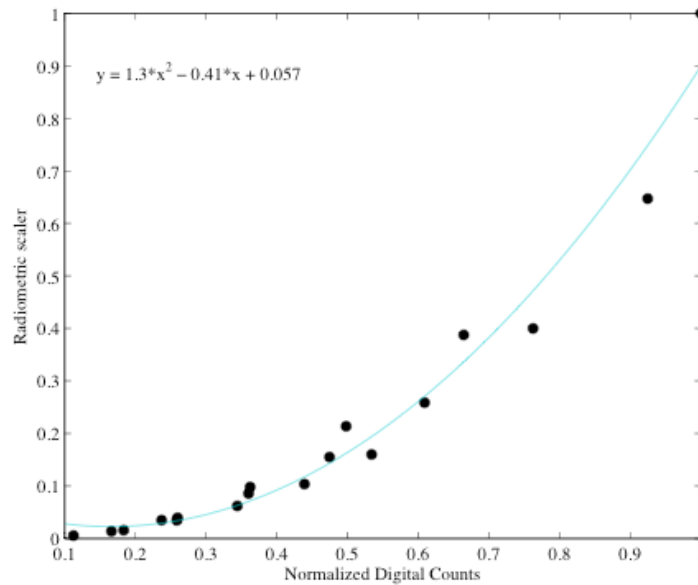


FIG. 2-5. Log luminance with the corresponding DC that have been shifted so they line up and then were fit with a curve that is a single channel's LUT.

These curves are then converted back to linear luminance and are used to linearize camera DC with respect to luminance. The matrix is generated for converting linear camera DC to XYZ values using non-linear optimization. The

starting values determined by solving an over determined system of linear equations to minimized sum squared error.

2.6 Display characterization

Modeling the display is important for psychophysics experiments that use the system to present stimuli. Accurate transformation from digital counts to colorimetric output of the display enables the stimulus to be the expected colorimetric value. Without characterization, unknown physical properties would be displayed leading to an inability to reach conclusions. Since the intention of psychophysics is to study the relationship between physical properties and perceptions, the results are dependent on having precise knowledge of the physical stimulus.

To characterize the color of a liquid crystal display (LCD) the Day et al. method [Day04]. The model consisted of three one-dimensional LUT and a 3x4 matrix transformation. The LUT describe the relationship between the signal used to drive the display and the radiometric output per channel. Measured data is used to define starting conditions that are then optimized to create the final model. Verification of this model is performed in preparation for subsequent psychophysics experiments. Comparisons of predicted and measured values of test data are with low color differences indicate a well-modeled system.

The Day et al. [Day04] method characterizes an LCD using a linear model and user control model [Berns97]. A linear model requires that the system be additive and scalable. The system is additive if the sum of the luminance of each

channel in isolation equals the luminance when all three are on simultaneously. Scalability exists when a single luminance curve of a primary can be multiplied by factors to reproduce the spectra at every drive level. Assuming additivity and scalability exist the radiance can be described as a linear combination of the radiance of each channel:

$$L_{\lambda,mix} = RL_{\lambda,r\max} + GL_{\lambda,g\max} + BL_{\lambda,b,\max} . \quad (2.5)$$

Which can be rewritten in terms of tristimulus values,

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_{r,\max} & X_{g,\max} & X_{b,\max} \\ Y_{r,\max} & Y_{g,\max} & Y_{b,\max} \\ Z_{r,\max} & Z_{g,\max} & Z_{b,\max} \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} . \quad (2.6)$$

where R , G and B are radiometric scalars of the channels. These values correspond to the user controls of the display and are characterized using a nonlinear relationship. For an LCD the OETF depends on the crystal construction, settings, and the internal LUT within the device. Due to the inherent nonlinearity the digital counts are transformed to radiometric scalars using three one-dimensional LUT. The linear and nonlinear portions are included in the Day et al. [Day04] method with the introduction of the of a flare correction that will be referred to as black correction. With the introduction of the black correction Eq. 2.7 is now,

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_{r,\max} - X_{k,\min} & X_{g,\max} - X_{k,\min} & X_{b,\max} - X_{k,\min} & X_{k,\min} \\ Y_{r,\max} - Y_{k,\min} & Y_{g,\max} - Y_{k,\min} & Y_{b,\max} - Y_{k,\min} & Y_{k,\min} \\ Z_{r,\max} - Z_{k,\min} & Z_{g,\max} - Z_{k,\min} & Z_{b,\max} - Z_{k,\min} & Z_{k,\min} \end{bmatrix} \begin{bmatrix} R \\ G \\ B \\ 1 \end{bmatrix} . \quad (2.7)$$

The LUT are unchanged because the radiometric scalars are unaffected by flare.

The tristimulus value matrix is optimized in a non-linear fashion to minimize the

color difference between the measured data and the model predicted data with each iteration updating the LUT. The resulting 3x4 transformation matrix and three one-dimensional LUT colorimetrically characterize the display. Evaluation of the model should be preformed on a test data set that was not used to generate the model. FIG. 2-6 is a flowchart of the characterization procedure laid out by Day et al.

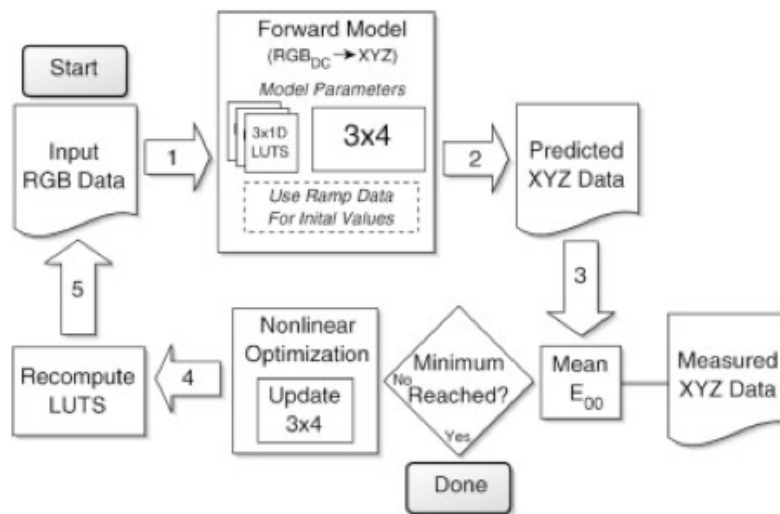


FIG. 2-6. Flowchart of characterization procedure used in Day et al. model [Day04].

3. Experimental Methods

3.1 Purpose

Overview

The research reported in this thesis contributes to the understanding of white point and chroma preferences of images that are cropped or placed close proximity to other images taken under different lighting conditions. A pair of psychophysical experiments designed to elicit information concerning these questions is described in the subsequent sections of this chapter. Experiment I looked at one image at a time and examined adjustment of white point, chroma or both. Experiment II looked at pairs of images and examined similar image adjustments on either one or both of the presented images. Within each experiment, the individual tasks were broken down into various parts distinguished by the controls available to the observer for image adjustment.

Previous work by the researcher and her advisor involved design of experiments that allowed real-time modification of white point and chroma for video. The current research, unlike those prior experiments, relied on presenting only still images but there was great benefit to the current experimental design from understanding gained through that prior experience. The experimental design for the previous research is also described later in this chapter.

Experiment I: Images in Isolation

The intent of Experiment I was two-fold. The first intention was to provide a baseline for how observers set preferred white point adjustments for images in isolation. This task is similar to past research into preferred reproduction.

The second intention of Experiment I was to learn about the relationship between cropping and preferred chroma. By comparing preferred chroma adjustments for cropped areas of images with the full images, an understanding would be built for how the apparent distance of an object from a camera changes people's expectation of chroma.

Cropped and uncropped images were mixed and presented in random order. Experiment I consisted of three parts. Part 1 offered adjustment for only white point. Part 2 allowed observers to only adjust chroma. And Part 3 allowed both adjustments simultaneously.

Experiment II: Images Adjusted Simultaneously

The intent of Experiment II was to derive data associated with the question of whether images taken under different illumination conditions and adjusted simultaneously differ from those adjusted in isolation. The results from this experiment were compared to those in Experiment I.

In Experiment II, two images were presented on the screen at the same time. For Part 1, only one of the two images was adjusted. In Part 2, both images were adjusted to create a cohesive page.

3.2 Design

Method of Adjustment

The psychophysical method used in these experiments was the method of adjustment. The precise range and threshold for both white point and chroma were unknown. This method is quick and easy to implement and is typically used when establishing a baseline in a new area of research. Decisions were limited to the ranges that would be attainable on the display for both white point and chroma.

Inspiration for User Interface

The inspiration for the experimental design in this research stems from earlier research looking at skin tone reproduction from web cameras used under various illumination conditions. In that experiment, observers were asked to adjust the hue and chroma to a preferred skin tone reproduction. The interface for the experiment was written in MATLAB R2007b (7.5.0.338) using PsychToolbox(3.0.8) [Brainard97]. Loading and processing time for each video clip was a considerable factor in the design. For that reason, preprocessing was performed and variables were saved into a MATLAB workspace file, where possible.

Real-time image processing was made possible by carefully reducing the number of online complex computations. Instead of performing computations on a per frame per pixel basis, computations were performed on the colormap matrix and the colormap was applied to each individual frame. A 16-bit colormap created from an example frame was used to convert frames to index color. Each time the

user modified a user control (changing chroma or hue of the video) a new colormap was derived. This was a fast process. Applying the map to subsequent images was nearly instantaneous. The newly color-modified frames were converted from index color image to a full 24-bits for display. Application of the colormap for conversion to true color takes place in less than a thirtieth of a second allowing for real-time playback.

For that experiment, input videos were assumed to be in the sRGB color space. This assumption allowed for a well-defined conversion to colorimetric values. Precise colorimetry of the input was not required for this experiment, since the goal was to understand the preferred colorimetry of the output.

Users were given hue and chroma controls in the GUI that were used to modify the sRGB values in CIELAB space. Input from the interface color controls were interpreted as changes in hue angle and chroma that were added onto the initial colormap in the CIELAB LCh color space. Display characterization via the Day et al. model enabled the conversion of the user manipulated colormap into known display colors [Day04]. Each frame was converted back to a true color image using the user manipulated colormap in real-time. FIG. 3-1 is a flowchart of the color workflow used to create, display and manipulate the videos.

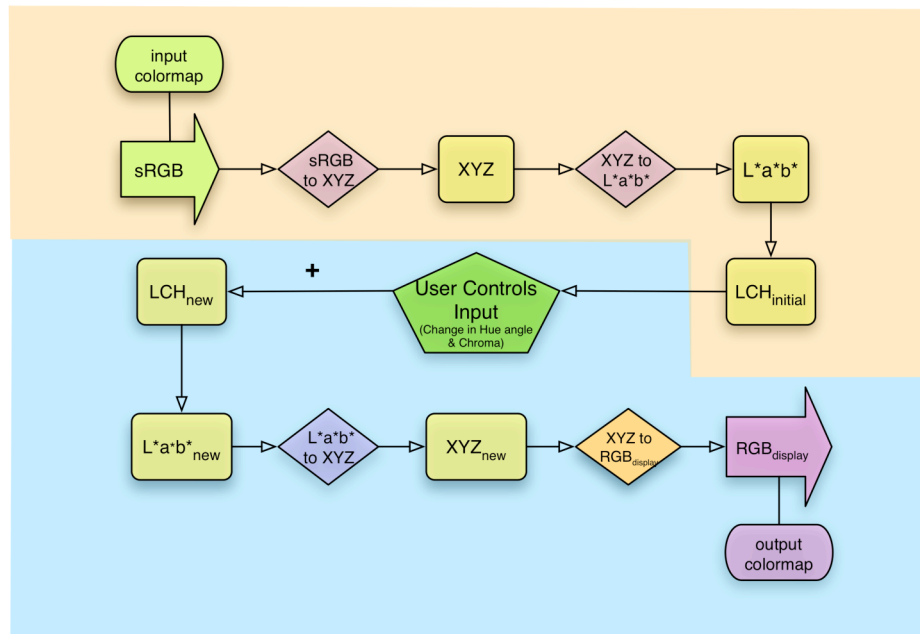


FIG. 3-1. Flowchart for inspiration experiment where the orange region is the processing that occurred prior to running the experiment and the blue region is the part that was preformed during the experiment.

The user interface from this experiment, shown in FIG. 3-2, was altered for this new research into white point and chroma adjustments.

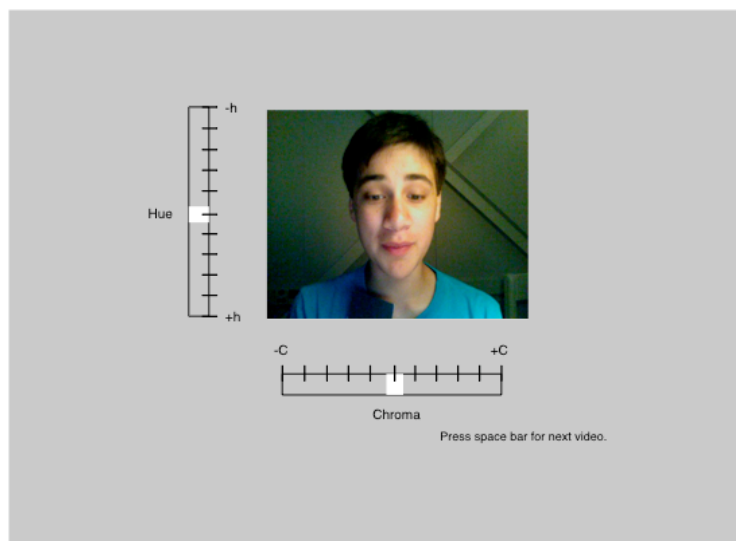


FIG. 3-2. Video experiments user interface.

The knowledge gained from this video-based experiment was applied to the new still-image experiments in this research. The use of colormaps made it again possible to quickly respond to observer settings. The experience gained from using PsychToolbox made the process of setting up the current experiments easier. The experimental design greatly benefited from this earlier work in skin tone reproduction.

3.3 Experiment Design

Experiment I: Images in Isolation

There were three parts to Experiment I as summarized in Fairchild's Levels

Fairchild [Fairchild05] described levels of reproduction based on Hunt's objectives for modern color imaging systems. Summaries of the levels are included in Table 2-II. Each level must satisfy the requirements of the previous level to define a reproduction at the succeeding level.

Table 2-II. All parts of Experiment I involved images presented one at a time. In Part 1 of Experiment I, data was acquired for white point preference. This provided a baseline useful for comparison with images whose white point was adjusted in the presence of another image taken under different illumination conditions. Part 2 of Experiment I examined the chroma adjustment. Part 3 of Experiment I included both white point and chroma controls. Parts 1 and 2 were meant to separate the interactions between the white point adjustment and the chroma adjustment, if any occurred.

Table 3-I. Experimental design for Experiment I summary include the parts and the corresponding controls provided to the observer.

Experiment I: Images in Isolation
Part 1: Adjust white point only
Part 2: Adjust chroma only
Part 3: Adjust white point and chroma

FIG. 3-3 is the user interface for Part 3 of Experiment I where both chroma and white point setting controls were present. For Parts 1 and 2, only the control adjustment was presented in the same locations as shown. The interface included a white border around the image to provide a reference white for the observers corresponding to the white point of the display.

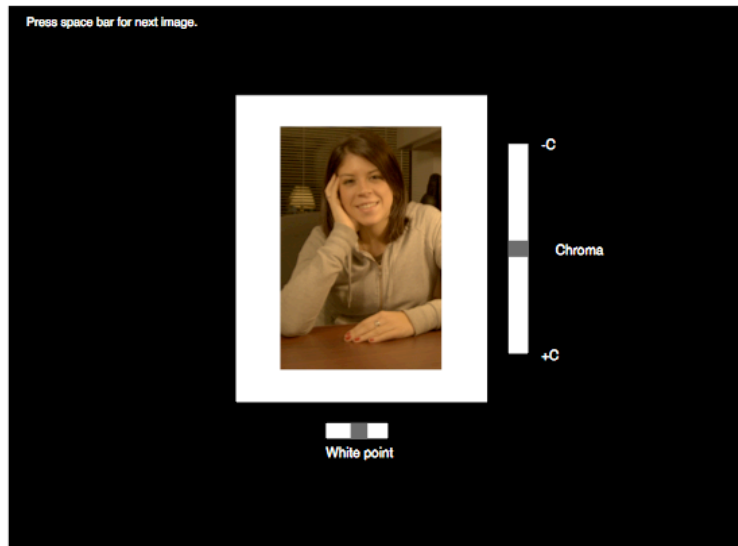


FIG. 3-3. Experiment I Part 3's user interface, which included both the white point and chroma user controls.

Experiment II: Multiple Images Presented Simultaneously

Experiment II had two parts which are summarized in

Table 3-II. Throughout this experiment, two images were presented to the observer at the same time, however only the white point user control was available to the observer. Part 1 of Experiment II had only one image in the pair adjusted at a time. Part 2 of Experiment II had controls for adjusting both images. Comparison of the results from Experiment II to those from Experiment I Part 1

examine whether or not there is a difference between images adjusted in isolation and those corrected in the presence of images taken under a different illuminant.

Table 3-II. Experimental design for Experiment II summary include the parts and the corresponding controls provided to the observer.

Experiment II: Images Adjusted Simultaneously
Part 1: Image on right white point adjusted
Part 2: Both images white points are adjusted

FIG. 3-4. Experiment II Part 2 user interface, which includes the white point controls for both the left and right hand images. Part 1 only one white point control was presented in the center of the image to control the image on the right.

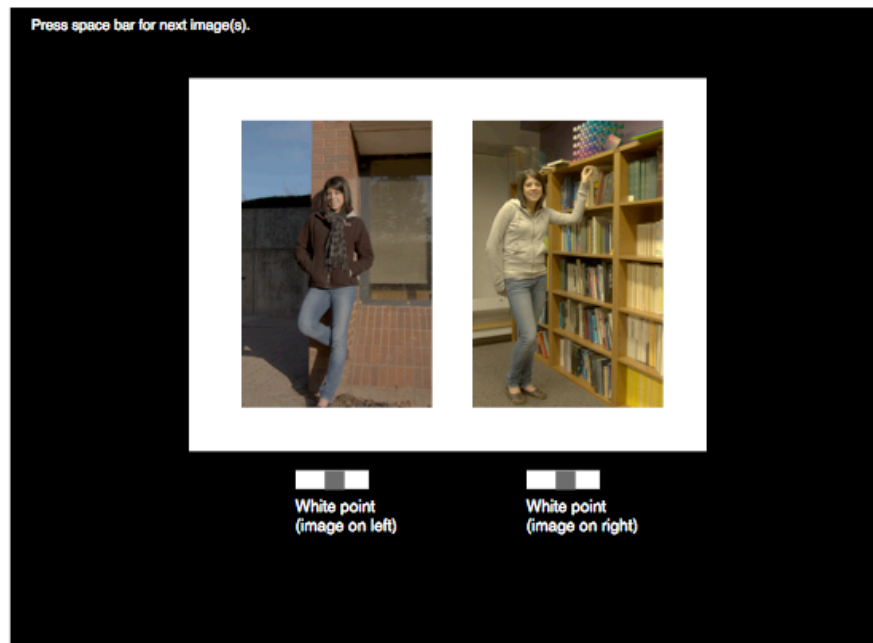


FIG. 3-4. Experiment II Part 2 user interface, which includes the white point controls for both the left and right hand images.

User Controls

The white point user control was designed so that the observer could pay attention to the image and not the control while making course or subtle changes to white point. Once observers clicked on the slider and held down the mouse they could move to the right or the left to change white point. Once the slider touched the edge of the slider bar, the white point would modify along the Planckian locus. It would continue to change until the edge of the slider was no longer touching the wall. Switching to the other wall caused the white point to shift in the opposite direction along the Planckian locus function. If the slider crossed the centerline of the control but did not hit a slider wall, adjustments paused. Once the mouse key was released, the slider snapped to the center position. This design ensured that the observers could not just place the slider in the same location for each image. Inherent in this type of control is variability in the intraobserver data. The chroma slider functioned based on direct difference in position requiring that the slider be long enough to provide the observer fine tune adjustment.

Changes in the white point occurred by incrementing the x chromaticity value starting at the measured white under the conditions that each photograph was taken. The y chromaticity value is a function of the x chromaticity value along the Planckian locus. The range of white points was restricted between 0.213 to 0.568 x chromaticity units with an increment size of 0.005. FIG. 3-5 graphically depicts the Planckian locus function used as a solid gray line and the range is restricted between the black dashed lines.

Measured white points for two of the three illumination conditions (triangles in FIG. 3-5) fall on this line, but fluorescent did not. It was important

that initial white point be included in possible resulting white points so that observers could potentially return to the initial state. Therefore the measured fluorescent white point was translated onto the curve. This translation was simply inputting the x chromaticity coordinate into the function to predict the corresponding chromaticity coordinate into the function to predict the corresponding y value.

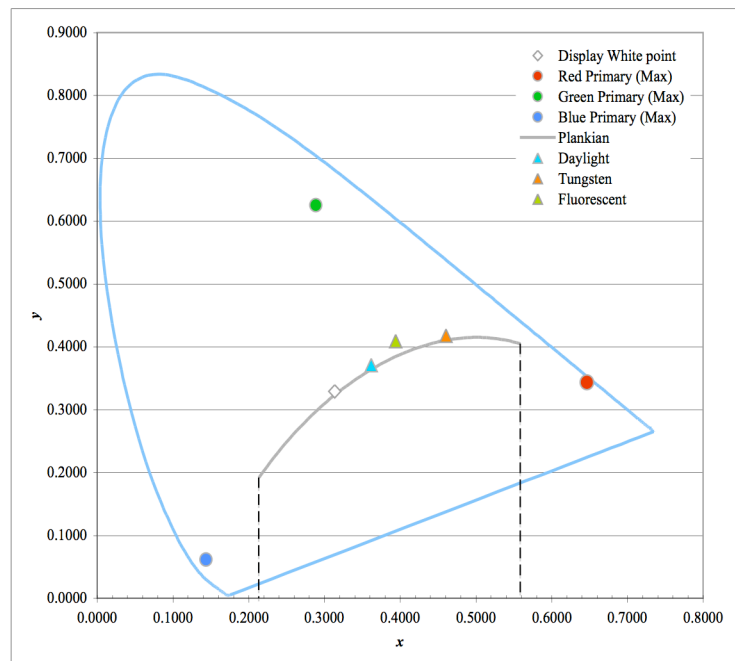


FIG. 3-5. Chromaticity plot that includes display primaries (circles) and white point (diamond) and the measured white points for each illuminant along with the Planckian locus function (gray curve) limited to the range (indicated by black dashed lines) used in the white point control.

The chroma slider functionality was similar to typical sliders in that as the observer moved the slider, the chroma changed by some amount based on the distance the slider was moved. The calculation of chroma was always based on an

initial difference of zero. The range was large enough to reach a neutral and a high chroma while remaining within the gamut of the display. A slider similar to the white point slider was not necessary because it was believed that observers would not intentionally set the chroma to the exact same spot since the initial chroma levels were not equal for each illumination condition. Between the long and zoom shots within a given illumination condition the initial chroma level were the same.

Image Processing Workflow

The color transformation workflow used in this research differed slightly from the one shown in FIG. 3-6 for the previous video-based experiment FIG. 3-1. There are several important distinctions. Instead of assuming the images were sRGB, the pixel colorimetry was defined by a characterization of the camera under the taking conditions. A normalization step was introduced to ensure that the luminance for all white points was considered the same. Dependence of the chroma calculation on the white point setting introduced new computations.

It was important to have the exact colorimetry that existed in the original scene in order to properly change the white points. Fully characterizing the camera and display system produced output that is referred to as a colorimetric reproduction. The specifics of the image capture process are provided in the following subsection.

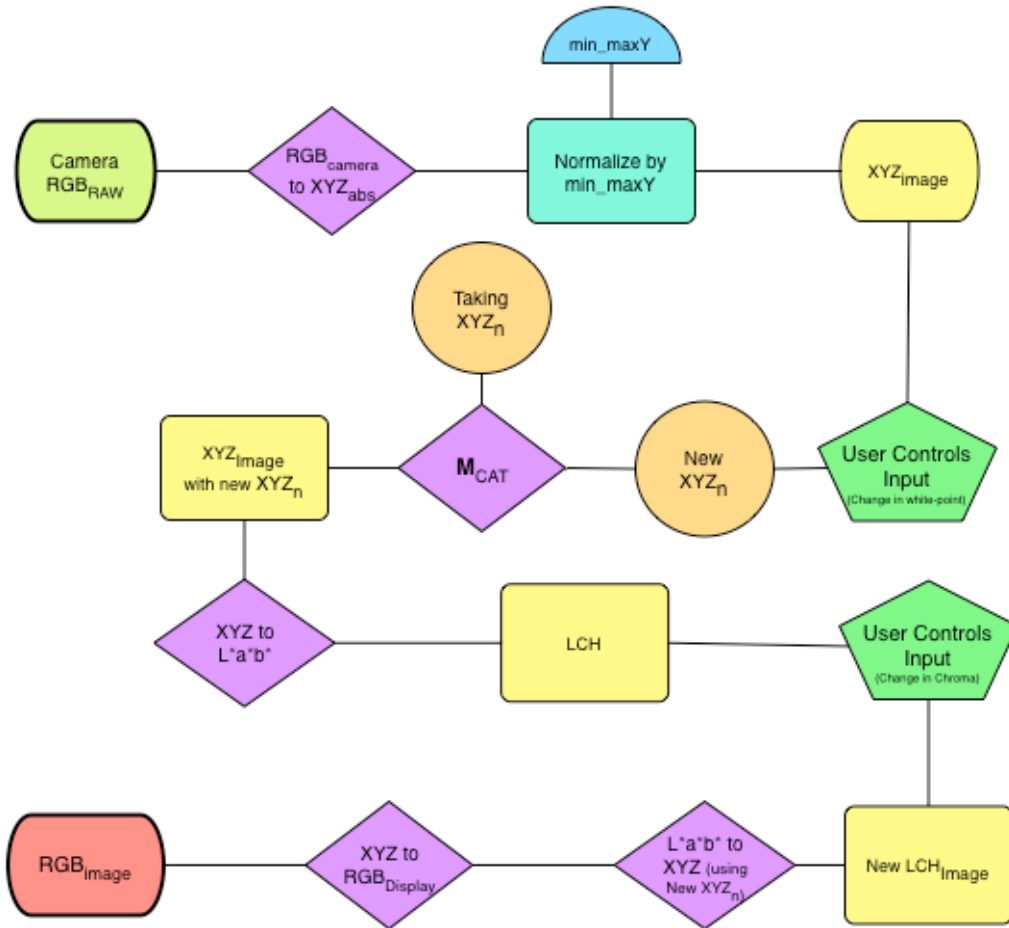


FIG. 3-6. Image processing workflow for this research.

A normalization step was introduced to ensure that the maximum luminance of the display did not change with a change in white point. The maximum luminance for the display is achieved at the white point of the display. Alternative white points naturally result in a decrease in maximum luminance available for a white. The maximum luminance available on the display was calculated for a sampling of possible white points within the white point range previously defined by the user controls. FIG. 3-7 contains the results of this analysis where the minimum of all maximum luminances available in that range is

0.34 (in relative tristimulus units) at an x chromaticity value of 0.213. Tristimulus values were normalized by this factor ensuring that maximum luminance remained constant with changes in white point.

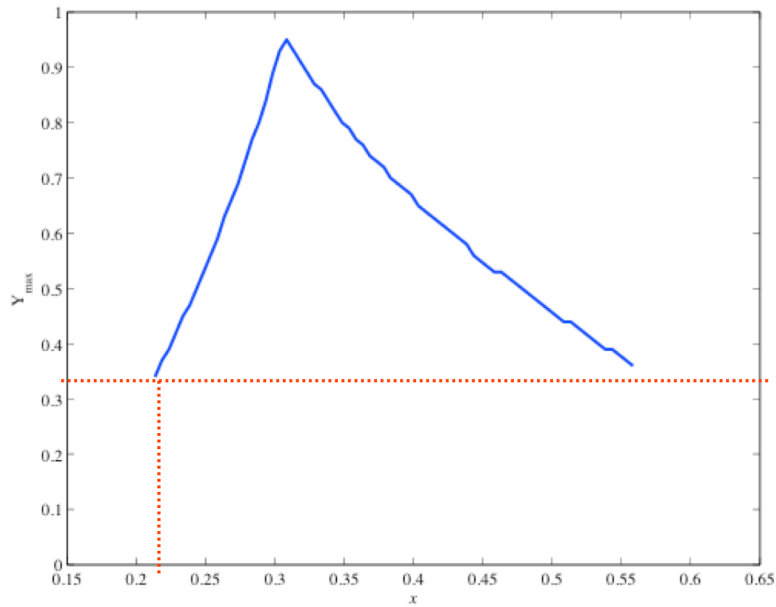


FIG. 3-7. Maximum luminance of the display for a given x chromaticity value (blue line) location of minimum is indicated by the red dashed lines.

Changing white points requires a way to chromatically adapt to the new white point. As previously discussed in the Background section, the von Kries model is commonly used in consumer cameras white balance algorithms and therefore was used here as well. This chromatic adaptation was performed prior to CIELAB values. Chroma calculations depend on the white point used to convert from tristimulus values to CIELAB. If only one control was adjusted then the other control values would be unchanged and skipped over in the transformation workflow.

User Instructions

The instructions for the Experiment I the single image experiment were:

You will be presented with several images and captions which are to be included in a photobook you are creating. Using the slider bar(s) provided adjust images to your preferred reproduction. There are three different control scenarios for which you will be able to adjust the images: white point alone, chroma alone and both white point and chroma together. When the preferred reproduction is reached press the space bar to proceed to the next image. Typically this experiment will take 15 to 20 minutes.

Instructions for the Experiment II the paired image experiment were:

You will be presented with several pairs of images and captions within a page layout, which are to be included in a photobook you are creating. In the first part of this experiment you will use a white point slider to adjust the image on the right (point to screen) so that the entire page layout including both images is your preferred page layout. In the second part of this experiment you will have two sliders one for each image (point to screen). Please adjust both images to your preferred page layout. When the preferred page layout is reached press the space bar to proceed to the layout. Typically this experiment will take 15 to 20 minutes.

3.4 Image Capture

The camera used to capture the images used in this experiment was a Nikon D40 with an 18 to 55 millimeter lens. An aperture setting of F/11 was used for all images. Image output format was raw (NEF for Nikon). Photoshop Camera Raw was used to demosaic the images. The raw converter was set in order to perform the minimum amount of processing and remained the same for all images.

Three illumination conditions were included in these experiments: daylight, fluorescent and tungsten. Measurements of the radiance of the illuminations were made using the PhotoResearch 650 spectroradiometer and

pressed polytetrafluoroethylene powder (Halon) as a perfect reflecting diffuser as a part of the characterization [Weidner81]. The color targets used in this characterization process were the Macbeth ColorChecker® DC as the training data and the Macbeth ColorChecker® as the testing data set.

As described in the background section, three exposures were taken of each target as well as the scene. The neutrals were used to create tone reproduction curves that include the full range of possible DC values. The final characterization LUTs are shown in FIG. 4-1 in the Results and Discussion section. The characterization was repeated under each lighting condition to ensure that the colorimetry was known. FIG. 3-8 contains the six images included in the experiment with two images per illumination condition. The top row contains the long shots and the bottom row contains the zoomed shots.

Creating Apparent Distance

The changes in the apparent distance were made in Photoshop CS3 Version 10.0.1 by cropping out a portion of each image. These images are referred to as the zoomed images. The original images will be referred to as the long shots. To create the zoomed images, the long shots were cropped to 399 by 600 pixels. Then the long shots were down sampled in Photoshop using the bicubic sampling method so that both the zoom and the long had the same pixel counts.



FIG. 3-8. Experiment images taken under three illumination conditions: (a) and (d) Daylight, (b) and (e) Fluorescent and (c) and (f) Tungsten. Images (a)-(c) are the long shots and (d)-(f) are the zoom shots.

3.5 Display Specifications

A 30" flat-panel Apple HD Cinema liquid crystal display was characterized and used in this experiment [Day04]. A Power Mac G5 controlled

the display through a DVI connection. It was set to maximum brightness using the buttons on the side of the display. The LCD is thin film transistor (TFT) active matrix display with 2560 x1600 pixel resolution. This monitor has 178° viewing angle and an antiglare coating. All surrounding illumination was turned off to provide a black background that allowed observers to completely adapt to the display's white point (D65).

3.5 Observer Population

All the observers had normal color vision. 29 observers participated in the single image experiment. The double image experiment included 28 observers. There were 21 males and 8 females ranging in age from 21 to 65 years old.

4. Analysis & Discussion

4.1 System Characterization

This research required two procedural elements to address the questions under study: 1) accurate knowledge of colorimetry from a scene and 2) accurate measurement of user-adjustments of that colorimetry on a screen. Camera characterization was repeated for each of the three illumination conditions. FIG. 4-1 illustrates sets of red, green and blue LUT for each condition derived using the technique described in Section 2.5. To convert to XYZ values, camera RGB digital counts are first passed through the appropriate LUT and converted to radiometric scalars. The scalars are then transformed using an optimized matrix to XYZ. The optimized matrices corresponding to the camera characterizations are shown in Eq. 4.1 for daylight, in Eq. 4.2 for fluorescent and in Eq. 4.3 for tungsten.

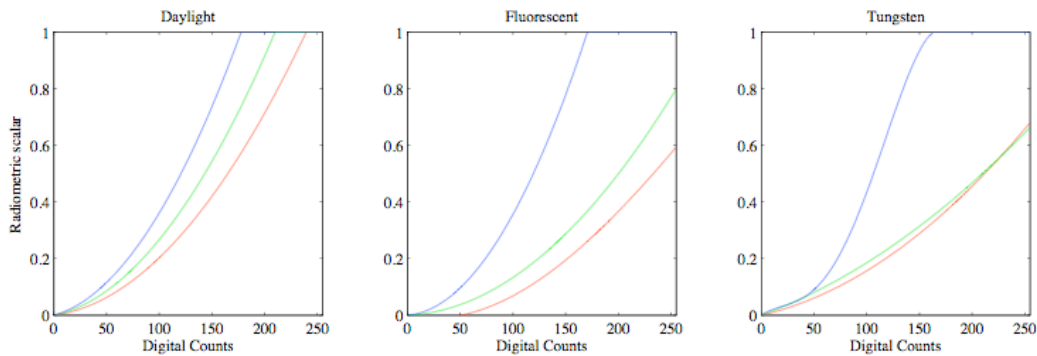


FIG. 4-1. Red, green and blue camera LUTs for each lighting condition. The line colors correspond to channels of the same color.

$$\begin{bmatrix} 0.658 & 0.277 & 0.09 \\ 0.274 & 0.728 & 0.00 \\ 0.00 & 0.00 & 0.634 \end{bmatrix} \quad (4.1)$$

$$\begin{bmatrix} 1.07 & 0.821 & 0.130 \\ 0.449 & 1.50 & 0.00 \\ 0.003 & 0.148 & 0.526 \end{bmatrix} \quad (4.2)$$

$$\begin{bmatrix} 0.789 & 0.1531 & 0.033 \\ 0.295 & 0.737 & 0.00 \\ 0.00 & 0.046 & 0.417 \end{bmatrix} \quad (4.3)$$

Verification of the performance of these characterizations are summarized in Table 4-1. Characterization of the daylight illumination condition was associated with the lowest mean CIEDE2000 units and the lowest standard deviation. Color error was found to predominately lie in the chroma direction. Illumination geometries for the scenes did not match each other or that of the color target. For the color patch target, colorimetric measurements were obtained with a geometry of 45/0. A chromatic adaptation transform was used to convert the measurement colorimetry to the scene colorimetry. The positioning of the targets within the scene was not identical between scenes. Attempts were made to minimize the changes, but differences were introduced between the positioning of training and test targets.

Table 4-1. Statistics for ColorChecker® verification of camera characterization under the three illumination conditions in terms of CIEDE2000.

Illumination Condition	CIEDE2000		Standard Deviation
	Mean	Maximum	
Daylight	3.07	5.99	1.51
Fluorescent	4.16	19.78	4.11
Tungsten	7.01	13.48	4.23

Good characterization of the display was. The additivity and scalability of the display were confirmed. The differences between the sums of the luminance of the three primaries compared to the luminance of the neutral ramp were less than 1%, which is well within the noise of the calculation (shown in FIG. 4-2). It is important to note that the sum of the three primaries includes three times the noise compared to the single measurement due to the fact that the noise is summed. FIG. 4-3 contains the chromaticities following optimized black correction of the three primary ramps. Subtracting off the flare the chromaticities of each primary map produces closer to a single coordinate. Noise prevents the primaries from truly plotting to a single chromaticity coordinate making it within the limits to have some slight variation. The display LUTs are shown in FIG. 4-4.

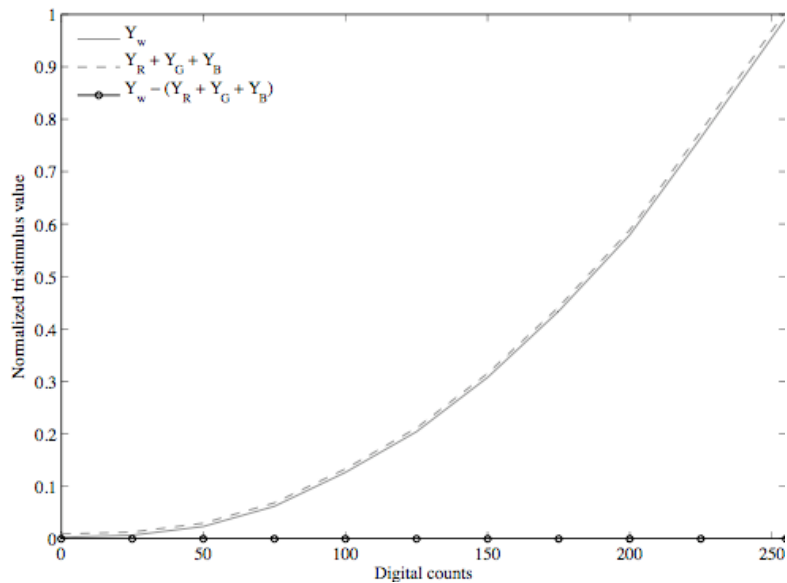


FIG. 4-2. Comparison of neutral ramp (Y_w) to the sum of each primary ramp ($Y_R + Y_G + Y_B$) to evaluate additivity.

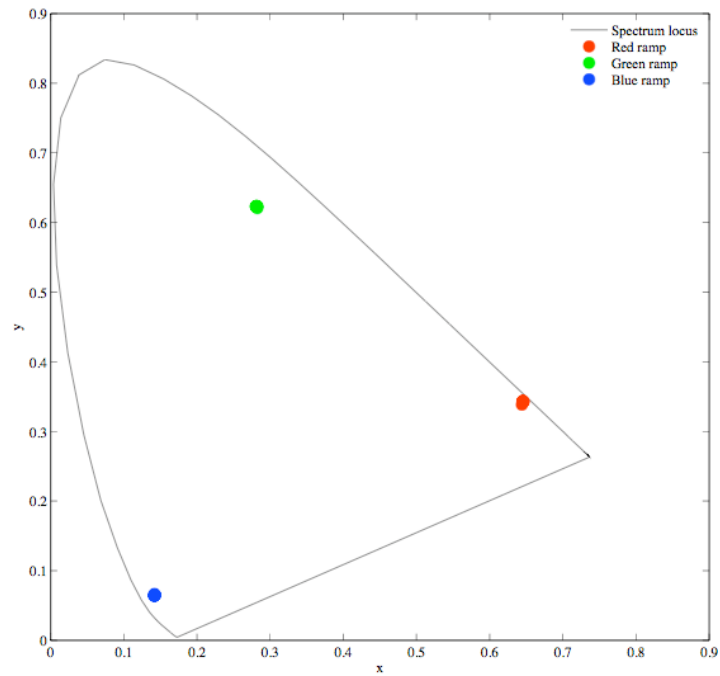


FIG. 4-3. Chromaticity plot of primary ramps excluding DC lower than 25 for the display after characterization optimization.

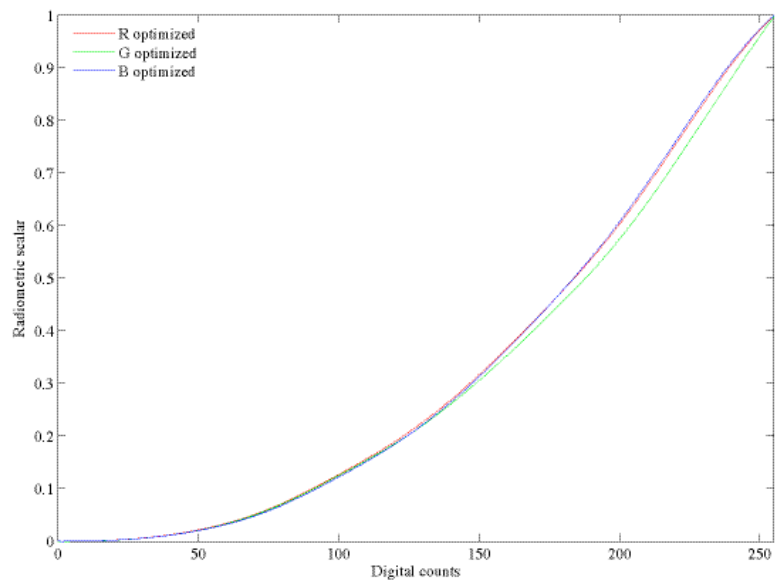


FIG. 4-4. Display LUTs between DC and radiometric scalars.

Verification of the display characterization resulted in small CIEDE2000 units for both mean and standard deviation summarized in Table 4-II.

Table 4-II. Statistics for verification dataset for display characterization in CIEDE2000.

	CIEDE2000		
	Mean	Maximum	Standard Deviation
Verification dataset	0.80	2.19	1.05

Acceptable system error characteristics were achieved in this verification process. This allowed for proper conversion of the camera digital counts to tristimulus values and then from tristimulus values to display RGBs.

4.2 Experiment I

Experiment I was designed to answer one of questions posed in this research, how chroma adjustments differ when apparent distance is changed, and to provide a baseline for the Experiment II. There were three parts to this experiment, in all the parts only a single image was presented (see FIG. 3-3). In Part 1, only the white point was adjusted, in Part 2 only the chroma was adjusted and in Part 3 both white point and chroma were adjust to made a preferred image. The starting images were fixed to have the same white point as the captured illumination and the same chroma as the original scene.

Chroma Adjustments

Results from Part 2 of this experiment were eliminated from the analysis based on observer opinions of the task. Observers felt that they could not achieve an acceptable color reproduction when only the chroma was adjusted. A trend was found that participants tended to adjust the chroma to a point that was lower than

the initial setting. Observers were clearly trying to eliminate the cast caused by the difference between the display white point and the taking illuminant colorimetry.

The relationship between the delta chroma adjustment for the long and the zoom image for each illumination condition from Part 3 are shown in FIG. 4-5. The line shown in each chart has a slope of one and represents where on the graph one would find the long and zoom images both adjusted the same or both left unadjusted. Points that lie close to the line only differed slightly. The majority of the results lay close to the line. Blue points are observations where zoom shots were adjusted to a higher chroma than long and magenta points were the opposite: long greater than zoom. For situations in which photographs were taken under daylight and tungsten lighting, there were more observations where zoom shots were adjusted to higher chroma compared to the long. It was observed that images taken under fluorescent had artifacts that were more pronounced in the zoomed image than the long shot. This may explain why the long shots for fluorescent taking environments were more frequently set to higher chroma than the zoom as these zoom-related artifacts were accentuated with increase in chroma.

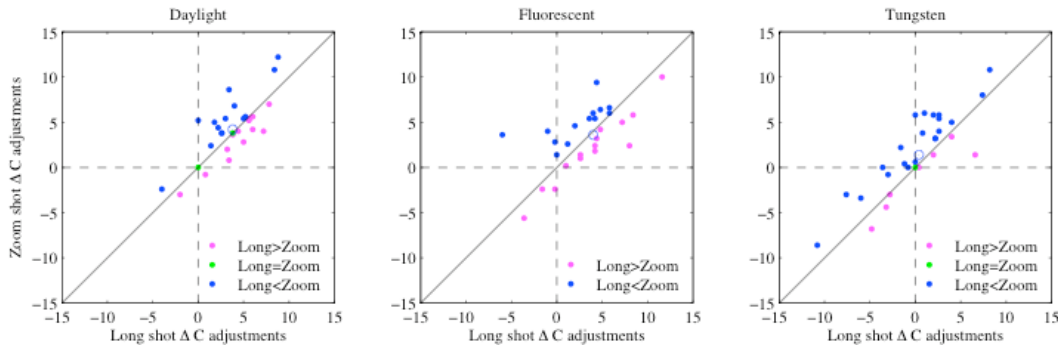


FIG. 4-5. Relationship between the ΔC^* adjustment for the long and the zoom image for each taking illumination condition.

Table 4-III contains the summary statistics of the chroma data from Part 3. Similar observations to those described for FIG. 4-5 were observed in these statistics. Nearly 60% of the population tended to adjust the chroma higher for the zoom shot than for long shot. These results were not statistically significant based on a maximum likelihood estimate of the probability with a 95% confidence interval. Equal differences in chroma were combined with the long-is-greater-than-zoom to determine the lack of statistical significance.

Table 4-III. Apparent distance percentage results for chroma for each illuminant.

	Long > Zoom	Long = Zoom	Long < Zoom
Daylight	38%	7%	55%
Fluorescent	59%	0%	41%
Tungsten	17%	3%	79%
Total	38%	3%	59%

White Point Adjustment

In Experiment I, the adjusted white point tended to move toward the display white point from its original position regardless of the illumination condition. This was an expected result, because people discount illumination in

the surround and the want their pictures to look like they remembered. Incorporation of white balancing algorithms within camera systems is done because consumers prefer the output images. Proper estimation of the white point of the taking illuminant allows a white balancing algorithm to eliminate a color cast caused by the scene's illumination.

The measured values in Table 4-IV are in CIEDE2000 units calculated from display white point to the initial white point of the taking illuminants. An average of the color differences of the adjusted white points are also listed. It should be noted that high variability existed in these results due to both observer preference and user controls. Comparing the difference between the measured and average adjusted CIEDE2000 values is shown in the third column of Table 4-IV. The CIEDE2000 values are all positive showing that on average, white points were adjusted toward the display white point. It is worth noting that after adjustment the images still maintain the same order of difference from the display white point as the scene measured white points.

Table 4-IV. CIEDE2000 between the display white point and the illuminant white point for both the scene measured and average adjusted white point.

	Measured	Avg.of adjusted	Meas. - Adj.
Daylight	17.0	8.2	8.8
Fluorescent	23.3	12.0	11.3
Tungsten	28.0	16.0	12.0

Statistical models were used to further analyze the white point data. Results are shown in Table 4-V. In isolation, preference adjustments of images taken under daylight were compared to adjustments of those taken under fluorescent and adjustments of those taken under tungsten. A p-value less than

0.05 are considered statistically significant. The white point adjustments for daylight images and fluorescent images were not statistically different from one another with a p-value of 0.43. This means observers do not prefer significantly different white points for these two illuminants. Photographs under daylight and those under tungsten illumination were found to have significantly different preference settings with a p-value of less than 0.001. Cropping did not result in a significance difference between these two illuminants.

Table 4-V. P-value statistics for Experiment I when the daylight image white point was adjusted in isolation compared to both fluorescent and tungsten images adjusted in isolation.

	Daylight vs. Fluorescent	Daylight vs. Tungsten
p-value	0.4306	< 0.001

4.3 Experiment II

Experiment II was designed to answer the question of whether the preferred white point for photographs viewed in the presence of another image taken under a different taking illuminant differs from the preferred white point for an image adjusted in isolation (the results from Experiment I Part 1). There were two parts to Experiment II. In both parts, a pair of images was presented side by side (see FIG. 3-4). Part 1 provided control for adjusting the white point of only the image on the right hand side. Part 2 provided the participant controls for adjusting the white point of both left and right hand images. The starting images were initialized to have the same white point as their respective taking illuminant.

Only the three long shots were used in Experiment II to limit the amount of time and redundant responses needed.

Similar trends to those found in the first experiment were also found when the images were presented simultaneously. As in Experiment I, the white point on average was adjusted toward the display's white point in all cases. In Experiment II the data showed that when adjusted in the presence of another image taken under another source illuminant, the white points were not adjusted as close to the display white point as they were adjusted in isolation (Experiment I, Part 1). Fifty-five percent of the population set their white point farther away from the display white point when there was a second image present taken under a different illuminant, compared to the results from Experiment I.

This effect was less prominent when both images could be adjusted to produce the preferred page, as was the case in Part 2 of Experiment II. This is illustrated in FIG. 4-6 for the adjustment of images captured under daylight. A majority of the adjusted white points from Experiment I Part 1 were all between the display white point and the measured taking illuminant as previously stated. The display and taking white points are shown as stars in FIG. 4-6. The graph on the left is for Experiment II, Part 1. The graph on the right is for Experiment II, Part 2. In Part 1 observers could only adjust the white point of one image. Higher variability in response white points occurred for Part 1 of the experiment resulting in a larger spread of the data. The standard deviations were also lower for Part 2 as shown in Table 4-VI. For Part 1 there is a definite difference in skew between daylight adjusted in isolation and daylight adjusted in the presence of fluorescent

and tungsten in FIG. 4-6. Daylight adjusted in isolation was closer to the display white point than the when adjusted in the presence of a picture taken under another illuminant. There was a visible trend associated with the taking illuminant of the second image. For both Parts 1 and 2, the tungsten image had more frequent occurrences of adjustments greater than the taking white point.

Table 4-VI. Standard deviation of x chromaticity data for the 28 observers when daylight was being adjusted.

Adjusting Daylight	Exp I Part 1	Fluorescent	Tungsten
Part 1	0.019	0.033	0.035
Part 2		0.023	0.027

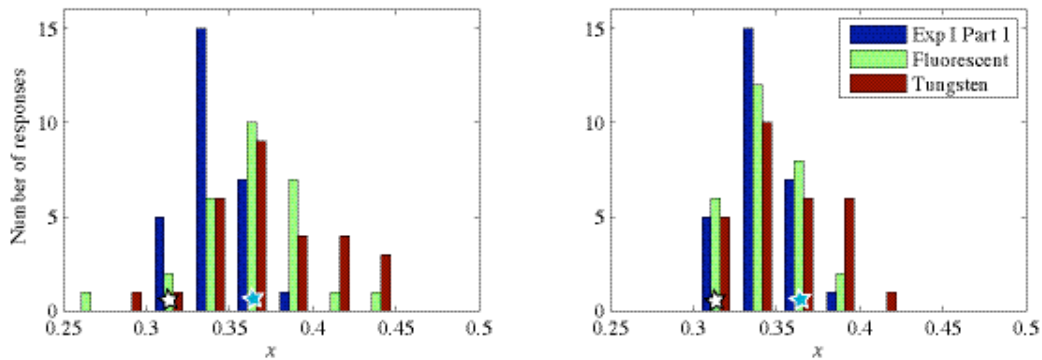


FIG. 4-6. Results from the 28 observers for Experiment II Part 1 on the left and Part 2 on the right compared to the results in blue for Experiment I Part 1 when daylight is adjusted. The white star is the display white point and the blue star is the measured taking illuminant.

The cumulative probability plots in Fig. 4-7 correspond to the data for adjusting images taken under daylight shown in FIG. 4-6 where the stars corresponding to the display white point are in white and the scene-measured illuminant has stars colored in blue. The left hand cumulative histogram shows clearly that observers made different choices when adjusting images to preferred

white point whether no image was present or an image taken under a different illuminant was present. Less clear, but following the same trend, we see in the right hand side of FIG. 4-6 that preference choices continue to be different even when both images can be adjusted as in Experiment II, Part 2.

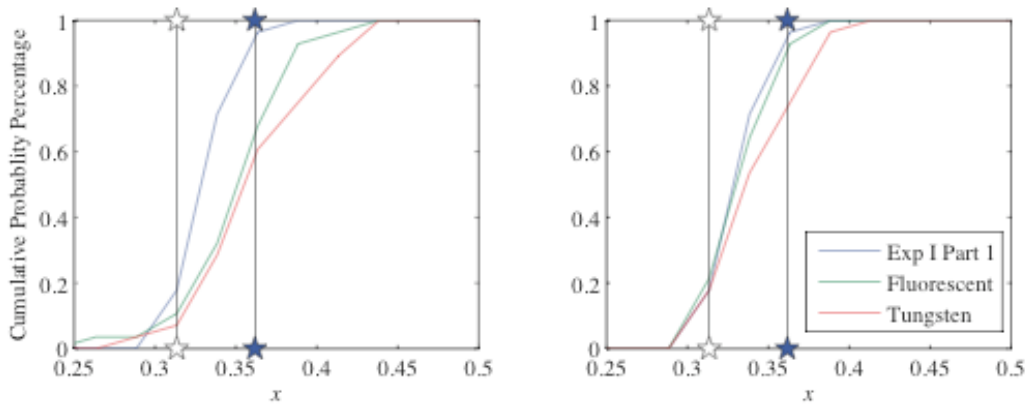


Fig. 4-7. Cumulative probability percentage plot for the histogram in FIG. 4-6. The white star corresponds to the display white point and the blue star to the scene measured daylight white point.

Adjusted white point results for the fluorescent image, in FIG. 4-8, were similar to those for the daylight taking image in that the majority of Experiment I Part 1 results were again between the display (white star) and taking illumination's white point (green star). When only a single image with a fluorescent white point could be adjusted as in Experiment II Part 1 (axis on the left) occurrences of white points shifted away from the display white point and toward the white point of the second image. The variability was also higher for Part 1 with the fluorescent image as shown in Table 4-X.

Table 4- VII. Standard deviation of x chromaticity data for the 28 observers when fluorescent was being adjusted.

Adjusting	Exp I Part 1	Daylight	Tungsten
Fluorescent			
Part 1		0.031	0.044
Part 2	0.034	0.022	0.031

In Experiment II Part 2, where controls are available for adjusting the white points for both images (axis on the right), a majority of the adjusted points were placed between the display white point and the taking illuminant. Fewer white points for the fluorescent source were adjusted further from the display white point and the taking illuminant than had been the case for daylight. There is a difference between fluorescent adjusted with daylight taking illuminant present and with fluorescent taking illuminant present. The image adjusted in the presence of a tungsten taking illuminant image generally was adjusted to a preferred white point further from the display white point than when adjusted in the presence of the image taken under daylight.

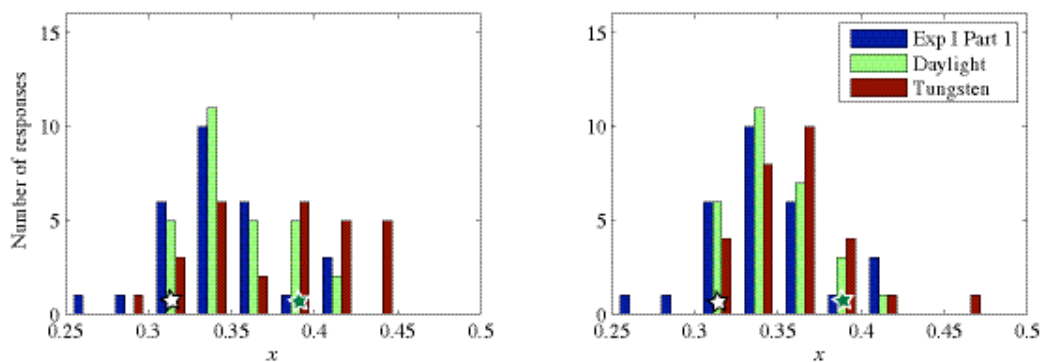


FIG. 4-8. Results from the 28 observers for Experiment II Part 1 on the left and Part 2 on the right compared to the results in blue for Experiment I Part 1 when adjusting fluorescent images. The white star is the display white point and the green star is the measured taking illuminant.

The cumulative probability plot that corresponds to the fluorescent image being adjusted is shown in FIG. 4-9. The display and scene-measured fluorescent white points are included as a white star and green star respectively. Once again there was a distinct difference between the three curves in both parts. This is especially evident for Part 2 on the right. This figure indicates that the preferred white point for a fluorescent image viewed in isolation is different from the preferred white point for the same image when viewed in the presence of an image taken under a different illumination condition, even when that image can be adjusted to a preferred white point.

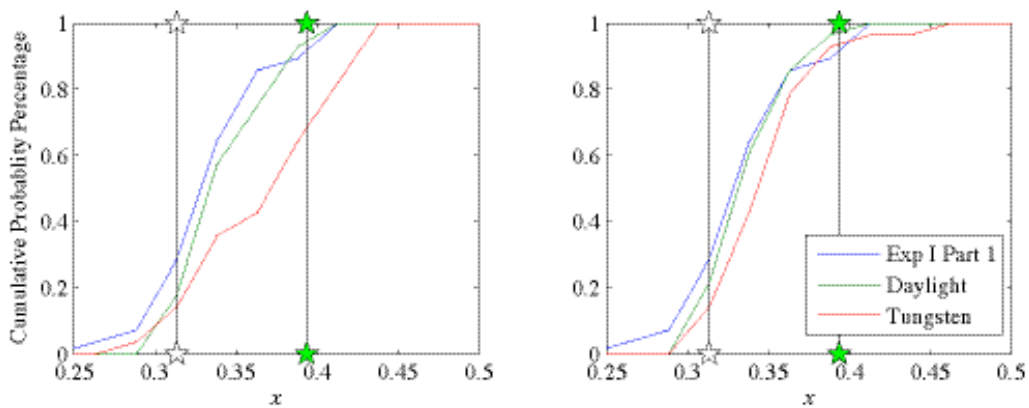


FIG. 4-9. Cumulative probability percentage plot for the histogram in FIG. 4-8. The white star corresponds to the display white point and the green star to the scene measured fluorescent white point.

FIG. 4-10 contains the results for when tungsten was adjusted. Compared to the previous results for the other two illuminants, the range of responses was much larger. The starting range between the scene measured and display white points was the largest for the tungsten scene, because the chromaticities for tungsten are larger than the other two illuminants. Observations covered the entire

range between the display and taking illuminant white points. For Part 1 (on the left) adjusted white points tended to lie half way between the display white point and the initial white point. All of the points in Part 1 were adjusted toward the display white point. The results from Part 2 (on the right) seem only slightly shifted toward the display white point compared to Part 1 results.

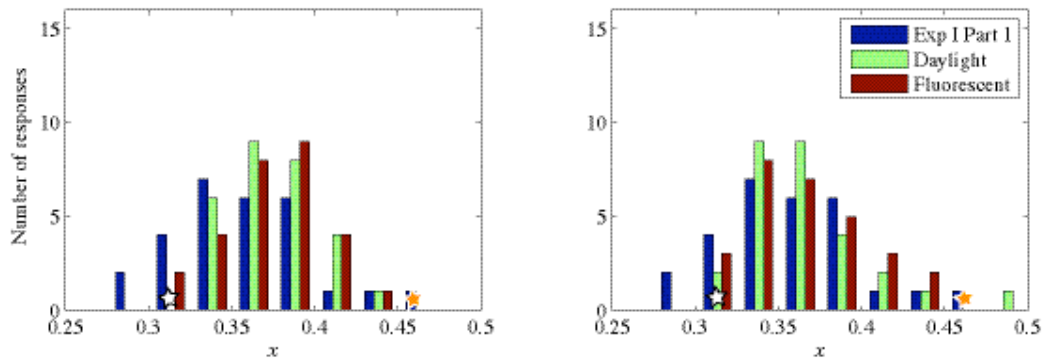


FIG. 4-10. Results from the 28 observers for Experiment II Part 1 on the left and Part 2 on the right compared to the results in orange for Experiment I Part 1 where tungsten was adjusted. The white star is the display white point and the orange star is the measured taking illuminant.

This observation was further supported in the cumulative probability plot in FIG. 4-11. A distinct difference exists for Part 1 and Part 2 between tungsten adjusted in isolation (Experiment I Part 1) and adjusted in the presence of either daylight or fluorescent balance images. This finding is based on the separation between the blue curve and the green and red curves, which represents the cumulative probability of x chromaticity in three different experimental situations. However a difference between tungsten adjusted in the presence of daylight or fluorescent is not obvious since the green and red curves are nearly overlapping.

This can be attributed to the fact that compared to the chromaticities of tungsten, those for daylight and fluorescent are considerably closer together.

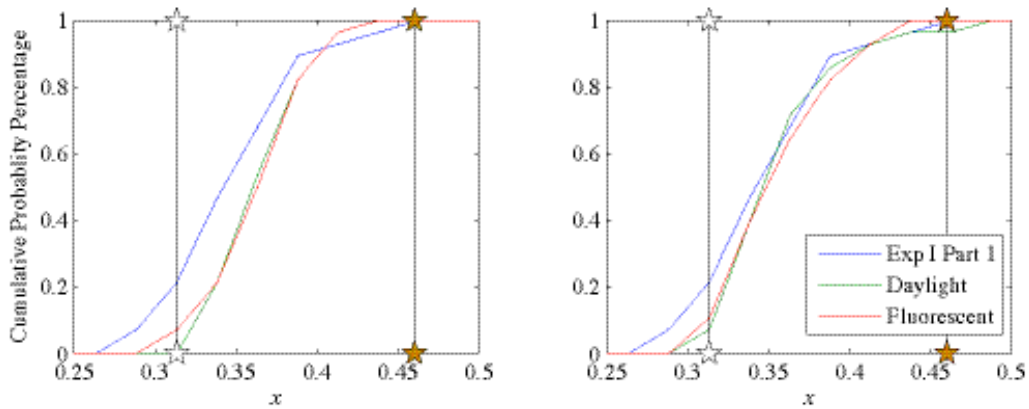


FIG. 4-11. Cumulative probability percentage plot for the histogram in FIG. 4-10. The white star corresponds to the display white point and the orange star to the scene measured tungsten white point.

For the tungsten image when adjusted in Part 2 (where both images were adjusted together to create the preferred page), the variability increased as shown in Table 4- VIII. This is different from the comparisons made when daylight and fluorescent images were adjusted. The lack of correspondence of these findings could be attributed differences in how the second image was adjusted. There as a large disparity between the tungsten white point and the daylight and fluorescent white points. This incongruity does not take away from the fact that there is a visible difference in the right hand part of FIG. 4-11 between the blue curve, (tungsten adjusted in isolation in Experiment I Part 1), and the green and red curves (where tungsten was adjusted in the presence of a daylight and fluorescent balanced images that were also adjusted to preferred white points).

Table 4- VIII. Standard deviation of x chromaticity data for the 28 observers when tungsten was being adjusted.

Adjusting Tungsten	Exp I Part 1	Daylight	Fluorescent
Part 1		0.026	0.030
Part 2	0.040	0.036	0.034

Further analysis was performed looking within observers comparing the results from images adjusted in isolation from Experiment I Part 1 to the two paired taking illuminant conditions from Experiment II when only one image was adjusted and when both images were adjusted. The results were sorted into 13 rank order categories from these comparisons. The results for adjusting the daylight image are summarized Table 4- IX where S is for Experiment I Part 1 (standalone) and F (fluorescent) and A (tungsten) represents the white points of the image in the second image. Bold lines box categories where the two white point adjustments were equal within a tolerance of one CIEDE2000 unit. The first condition in the row was adjusted to the highest color difference from the display white point. The last letter in the category was adjusted to closest to the display white point. Numbers of observations per category were summarized per part.

Table 4-IX. Ordered categorical results for the daylight image looking at the CIEDE2000 distance of the adjusted white point from the display white point. S (standalone) is from Experiment 1, Part 1 where the image was adjusted in isolation. F (fluorescent) and A (tungsten) are from Experiment 2 where the image was adjusted in the presence of the associated second image. Part 1 only allowed the first image to be adjusted. Part 2 allowed both image white points to be adjusted. Leftmost category is highest adjusted color difference from display white point, rightmost category is lowest adjusted color difference from display white point and bold line boxes represent multiple categories adjusted to preferred values that differ to within one CIEDE2000 unit.

Category	Part 1	Part 2
S F A	1	2
S A F	0	2
A S F	2	1
A F S	12	6
F S A	1	0
F A S	0	2
S F A	1	1
A S F	1	2
S F A	0	2
S F A	2	1
F A S	7	7
S A F	1	2
F S A	0	0
Total obs.	28	28

The category with the maximum number of responses for Part 1 was AFS where the daylight image paired with tungsten had the largest color difference from display white followed by the daylight image adjusted in the presence of the fluorescent image and then daylight adjusted in isolation in Experiment I Part 1. The category for Part 1 with the second largest was F=AS, which is similar to AFS except the difference between daylight adjusted with fluorescent present and daylight adjusted with tungsten present have less than one CIEDE2000 unit

difference between them. Similar results were found for Part 2 except with more variability in the distribution of the observations.

When an image captured under daylight was adjusted in isolation, the white point was shifted closer to the display white point than when daylight was adjusted with another image on the screen that was taken under a different illuminant. The strongest effect was exhibited in Experiment II Part 1 when daylight was adjusted in the presence of an image with a tungsten white point. Daylight was adjusted further from the display white point than when adjusted in isolation or in the presence of a fluorescent image. Based on previous research into color appearance and the effect of surround, these results support the notion that image appearance of a first image is impacted by the presence of a second image in its background. The trend found here was that subjects shifted the first image's preferred white point closer to the second image's taking white point.

Table 4-X summarizes the results for adjusting fluorescent images. Similar to daylight, the category with the most frequent responses for both parts was ADS. Fluorescent adjusted in the presence of a tungsten image pulled the fluorescent white point further from the display white point. For Part 2, D=AS was the category with the second most observations with seven observations. This category is similar to ADS except that there is no difference (within a one CIEDE2000 unit tolerance) between when the fluorescent image was adjusted with daylight and tungsten in the surround.

Table 4-X. Ordered categorical results for the fluorescent image looking at the CIEDE2000 distance of the adjusted white point from the display white point. S (standalone) is from Experiment 1, Part 1 where the image was adjusted in isolation. D (daylight) and A (tungsten) are from Experiment 2 where the image was adjusted in the presence of the associated second image. Part 1 only allowed the first image to be adjusted. Part 2 allowed both image white points to be adjusted. Leftmost category is highest adjusted color difference from display white point, rightmost category is lowest adjusted color difference from display white point and bold line boxes represent multiple categories adjusted to preferred values that differ to within one CIEDE2000 unit.

Category	Part 1	Part 2
S D A	1	5
S A D	1	1
A S D	0	1
A D S	11	7
D S A	0	0
D A S	2	1
S D A	2	2
A S D	3	0
S D A	1	1
S D A	1	1
D A S	3	6
S A D	3	3
D S A	0	0
Total obs.	28	28

Table 4-XI summarizes the rank order results where tungsten was adjusted. FDS had the highest frequency of responses for Part 1 where the image captured under the tungsten taking illuminant was adjusted in the presence of a fluorescent image was the furthest from the display white point tungsten. The tungsten image adjusted in isolation, Experiment I Part1 had the smallest color difference from the display white point. The results for Part 2 were highly variable as there was no category clearly in the lead. D=FS was most frequent for Part 2, but it is clear from the variability that a defining difference between

tungsten adjusted with daylight and fluorescent was not found. Combining the observations for categories where the tungsten adjusted in isolation, Experiment I Part 1, had the smallest color difference the result are 15 observations for Part 1 and 10 observations for Part 2.

Table 4-XI. Ordered categorical results for the tungsten image looking at the CIEDE2000 distance of the adjusted white point from the display white point. S (standalone) is from Experiment 1, Part 1 where the image was adjusted in isolation. F (fluorescent) and D (daylight) are from Experiment 2 where the image was adjusted in the presence of the associated second image. Part 1 only allowed the first image to be adjusted. Part 2 allowed both image white points to be adjusted. Leftmost category is highest adjusted color difference from display white point, rightmost category is lowest adjusted color difference from display white point and bold line boxes represent multiple categories adjusted to preferred values that differ to within one CIEDE2000 unit.

Category	Part 1	Part 2
S D F	3	3
S F D	2	2
F S D	1	0
F D S	7	4
D S F	1	2
D F S	4	1
S D F	0	3
F S D	2	1
S D F	2	1
S D F	1	4
D F S	4	5
S F D	0	0
D S F	1	2
Total obs.	28	28

Table 4-XII contains the adjustment conditions causing the smallest CIEDE2000 adjustment difference from the display white point. The symbols in the first column relate to those in previous three tables (Table 4-IX, Table 4-X and Table 4-XI) with the gray squares representing all the possible combinations

that end in the same symbol. The majority of the observations were included in categories that ended with the image in isolation (S) for all three-illumination conditions. This means that a majority of the images adjusted in isolation were preferred with white points closer to the display white point than for those adjusted with a second image present. This supports our hypothesis that preferred white point differs when an image is viewed in isolation with respect to when the image is viewed with a second one present.

Table 4- XII. Combined categorical results taken from Table XII, Table XII and Table XIV according to the smallest CIEDE2000 adjustment difference from display white where gray boxes are all possible symbols.

















White point of other image being presented simultaneously	White point of image being adjusted		
	Daylight	Fluorescent	Tungsten
 S	68%	46%	50%
 D	-	39%	25%
 F	14%	-	21%
 A	14%	14%	-
Ties	4%	1%	4%
Total	100%	100%	100%

Table 4- XIII contains the adjustment conditions causing the largest CIEDE2000 difference from display white. The symbols are the same in this table as they were in Table 4-XII. Daylight and fluorescent images adjusted in the presence of a tungsten image and the tungsten image adjusted in the presence of a daylight image had the largest CIEDE2000. When the white point of the image being adjusted was dissimilar to the second image, the white points were adjusted

further from display white. Tungsten images adjusted in the presence of daylight or fluorescent images are adjusted to similar white points. This further supports the observations that there was no statistical difference for adjusting tungsten in the presence of either daylight or fluorescent images. It is clear, though, that the presence of second images with different taking white points do cause a preference shift in preferred white points away from that preferred when the image is viewed by itself on the display.

Table 4-XIII. Combined categorical results taken from Table XII, Table XII and Table XIV according to the largest CIEDE2000 adjustment difference from display white where gray boxes are all possible symbols.

White point of other image being presented simultaneously	White point of image being adjusted		
	Daylight	Fluorescent	Tungsten
  	14%	25%	21%
  	-	14%	43%
  	29%	-	32%
  	54%	61%	-
Ties	3%	0%	4%
Total	100%	100%	100%

Images presented in isolation and simultaneously both tended to maintain the same order as their initial white points, shown in Table 4-XIV. Based on the tests concerning a population proportion, the percentage required for significance was 59% [Devore94]. The close proximity of daylight and fluorescent white points caused fewer occurrences of maintaining order when comparing these two. This was true for both isolation and simultaneous presentation of images. When

initial white points are further apart, then the order is more clearly preserved. A higher percentage of observations were observed for tungsten with daylight compared to tungsten with fluorescent.

The largest difference in initial white point was between daylight and tungsten, which influenced the percentage of observations that maintained order to be the highest at 75%. Daylight and fluorescent had the smallest difference leading to 50% of the observations maintaining their order.

Table 4-XIV. Percentage of observer responses maintaining relative order of adjusted white points for Isolation presentation and Simultaneous presentation and percentage of observer responses increasing the average white point distance from the display white point when comparing Isolation presentation to Simultaneous presentation.

	Maintain order		Distance increase in simultaneous case
	Isolation	Simultaneous	
Population	64%	61%	60%
Daylight vs Fluorescent	50%	43%	57%
Daylight vs Tungsten	75%	75%	54%
Fluorescent vs Tungsten	68%	64%	68%

Statistical models were used to determine if the difference between the measurements of images in isolation and the paired adjusted images from Part 1 were statistically significant. For all cases, a significant difference between the images in isolation and those adjusted with another image in the surround were found. The p-values are summarized in Table 4-XV where values less than 0.05 are considered statistically significant.

Table 4-XV. Experiment II Part 1 p-values for each illuminant.

White point being adjusted	Surround 1	Surround 2
Daylight	Fluorescent 0.0016	Tungsten 0.0015
Fluorescent	Daylight 0.0249	Tungsten 0.0441
Tungsten	Daylight 0.0079	Fluorescent 0.0105

From the trends shown in the tables, adjusting images in the presence of an image with a different white point does affect the adjusted white points. Preferred white points were chosen somewhere between where they started and the display white point. This enabled the image to differentiate itself from the display illumination and the other displayed image. Differences in taking white points between the two images do appear to influence the amount that the white points were adjusted to their preferred state. As expected, and based on current image appearance models, an image in the background of a first image will affect how the white point is adjusted. Even when both images were adjusted, a difference continued to exist between adjustments performed in isolation and those performed in the presence of an image taken under a different illuminant.

5. Conclusions

This research is a first look at two important questions: how does cropping impact chroma preference and how do multiple images on a page impact color correction preferences when each was taken under a different illuminant. Information gathered from this investigation and future ones should lead to improvements in image appearance models. Incorporating these improvements into algorithms for processing personalized consumer documents such as photo books will increase the quality of resulting output.

Experiment I addressed the impact of cropping on chroma preference and also established a baseline for color correction of the test images in isolation. Overall, the results indicated that observers did prefer higher chroma for cropped images versus preference for uncropped images. Cropping the images caused an increase in relative area of colors within the scene. However, Experiment I limited in the amount of cropping and the type of images included. Further research into this particular question should be done based on the trends established from this experiment's results. These trends were indicative of the zoom image being adjusted to a higher chroma. A cropped image that is made to be the same size as an uncropped image provides the appearance, to an observer, that the subject of the image is closer to the camera. Future work should include a wider range of apparent distances with a larger variety of image content.

In Experiment I, a baseline was also established for white point adjustments in isolation. On average, white points were adjusted toward the white point of the display regardless of the taking illuminant. Adjustments were made in such a way that the orders of distances of the taking illuminants from the display white were maintained after they were adjusted. The blue dots in FIG. 5-1, 5-2 and 5-3 represent the average of the results from the images adjusted in isolation for daylight, fluorescent and tungsten, respectively. The black diamond is the original taking illuminant. Consumer imaging manufactures have typically applied white balancing to images but usually it is used to remove original taking illuminant bias. Observer variability was high based on the psychophysical method.

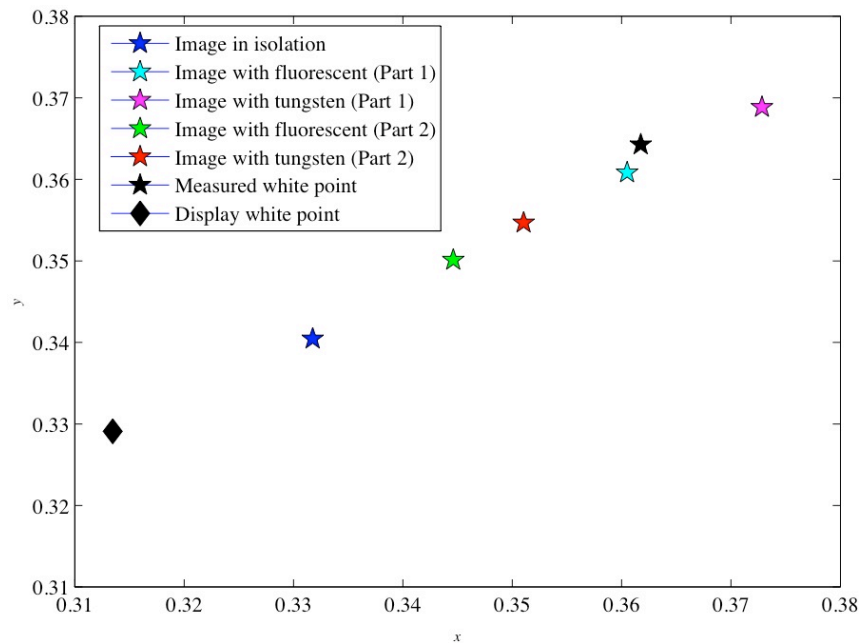


FIG. 5-1. Average results for daylight-captured image for the white point adjustment.

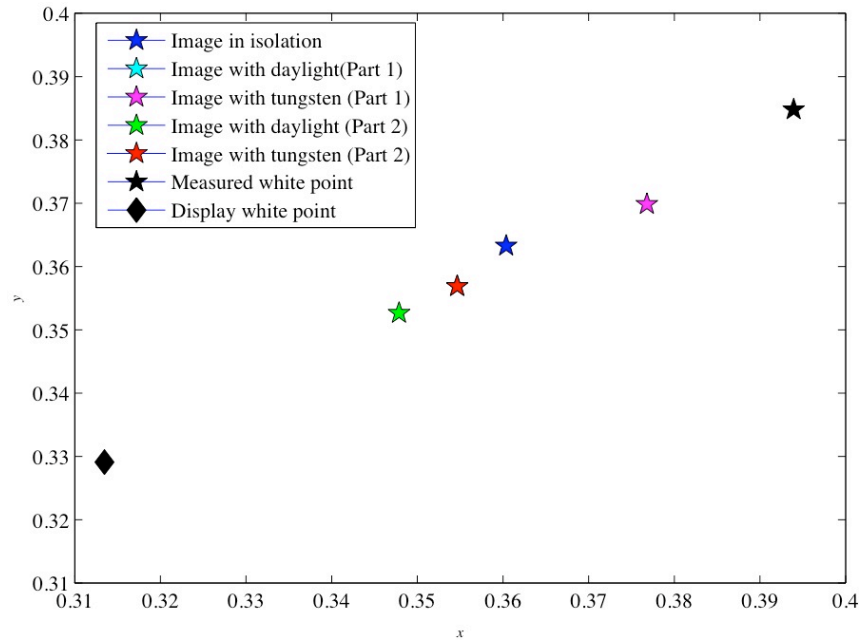


FIG. 5-2. Average results for fluorescent captured image for the white point adjustment.

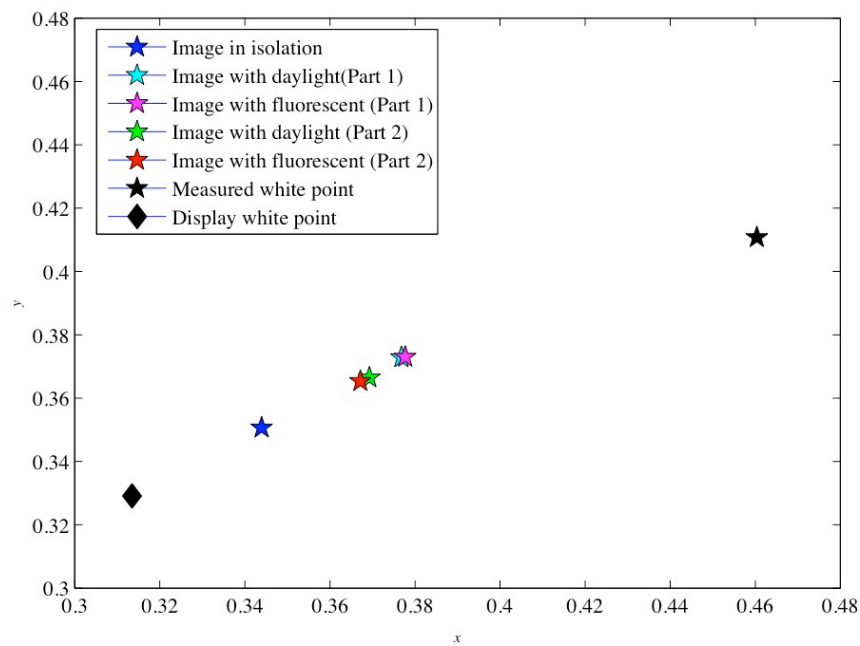


FIG. 5-3. Average results for tungsten-captured image for the white point adjustment.

The major finding from Experiment II Part 1 was that having a second image in the presence of the first image does affect the white point adjustment.

White points were shifted, again, toward the display's white point, but there was a significant difference between the results from the adjustment on the image in isolation (Experiment I Part 1) and with a second image present (Experiment II Part 1). The cyan and magenta dots in FIG. 5-1, 5-2 and 5-3 are the average chromaticity results for the two surround conditions for Experiment II Part 1. Generally the cyan and magenta points were the furthest away from the display white point. Tungsten taking illuminant had the largest affect on the image it was paired with by pulling the white point further from the display white point. Large differences in white points between the paired images on the same page were found to cause the white point to move further from the display white point.

Experiment II Part 2, where both the first and the second images were adjusted by the participants, exhibited similar results to Part 1, but to a lesser degree. Both images were adjusted toward the white point of the display causing the difference between the images to decrease. These differences were statistically significant. The tungsten taking condition, again, had a larger affect on the image it was paired with, compared to daylight and fluorescent. A significant difference in chromaticity was found to exist between Experiment II Part 2 and Experiment I Part 1. As shown in FIG. 5-1, 5-2 and 5-3 where the difference between Experiment II Part 2 and Experiment II Part 1 is depicted by the red and green dots and the cyan and magenta dots respectively. The red and green dots are closer to the display white point than the cyan and magenta dots from Experiment II Part 1.

In summary the main finding of this research was that there is a difference between images adjust in isolation and those adjusted in the presence of an image with a different taking illumination. FIG 5.4 summarizes these findings in the arrow pointing from the preferred area for images adjusted in isolation (blue cloud) and the area preferred for an image adjusted in the presence of another image with a differing white point (red cloud). This figure also depicts he preferred white points are shifted toward the display white point and away from the original white points (the magenta and cyan stars).

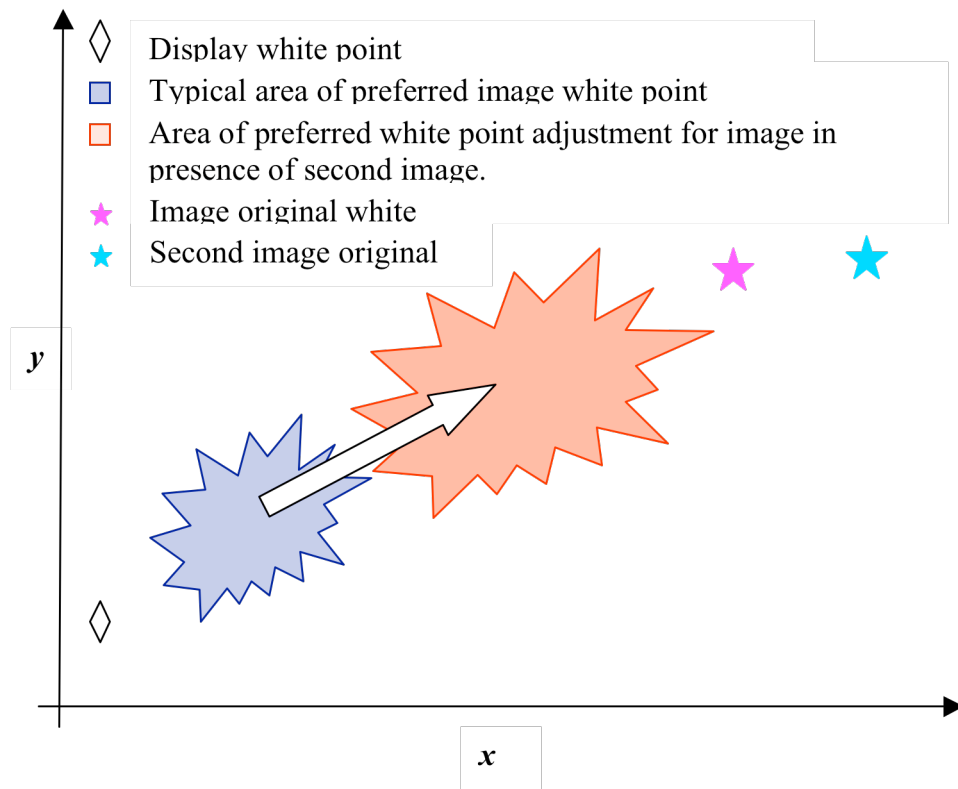


FIG. 5-4. Depiction of difference between image adjusted in isolation, the blue cloud, and images adjusted simultaneously in the presence of image 2, the red cloud. The magenta and cyan stars are the original white points for the image and the second image.

Future Works

In order to further address questions regarding the relationship between chroma and apparent distance, future studies should include an increased number of images with various image contents. The image database should include landscapes, as well as subjects at various apparent distances. These additions would verify the trends found in this research.

Future work to improve upon the white point adjustment findings should include comparing the current results to those predicted by current image appearance models for images presented simultaneously. If these models do not account for the trends found in this research, then modifications to these models would be made. Including more images with different content would expand the understanding of how white balance preference is impacted by other images taken under different conditions and should make applicability of these results more universal.

References

- [Bartleson59] Bartleson C. Some observations on the reproduction of flesh colors. *Photographic Science and Engineering*, **3**(3), 114-117 (1959).
- [Bartleson61] Bartleson C. Color in memory in relation to photographic reproduction. *Photographic Science and Engineering*, **5**(6), 327-331 (1961).
- [Berns97] Berns R. A generic approach to color modeling. *Color Research and Application*, **22**(5), 318-325 (1997).
- [Bodrogi01] Bodrogi P & Tarczali T. Colour memory for various sky, skin, and plant colours: effect of the image context. *Color Res. Appl.* **26**(4), 287-289 (2001).
- [Brainard97] Brainard, D. H. The Psychophysics Toolbox, *Spatial Vision* 10:433-436 (1997).
- [Braun07] Braun K. Method for determining an adjustment amount to an input chroma. US Patent 7196735: 27 March 2007.
- [Buck48] Buck G and Froelich H. Color characterization of human complexions. *Illuminating Engineering*, **43**: 27-49 (1948).
- [deRidder96] de Ridder H. Naturalness and image quality: saturation and lightness variation in color images of natural scenes. *J. Imaging Sci. Tech.* **40**(6), 487-493 (1996).
- [Day04] Day E, Taplin L & Berns R. Colorimetric characterization of a computer-controlled liquid crystal display. *Color Research and Application*, **29**(5), 365-373 (2004).
- [Devore95] Devore J. Probability and statistics for engineering and the sciences (Ed. 4th). Duxbury Press: Pacific Grove. (1995).
- [Fairchild05] Fairchild M. Color appearance models (2nd ed). John Wiley & Sons Ltd. Chichester, England: (2005).
- [Fedorovskaya97] Fedorovskaya E, de Ridder H & Blommaert F. Chroma variations and perceived quality of color images of natural scenes. *Color Res. Appl.* **22**(2):96-110 (1997).
- [Fernandez02] Fernandez S & Fairchild M. Observer preferences and cultural differences in color reproduction of scenic images. *IS&T/SID 10th Color Imaging Conference*, Scottsdale, 66-72 (2002).
- [Fernandez05] Fernandez S, Fairchild M & Braun K. Analysis of observer and cultural variability while generating “preferred” color reproductions of pictorial images. *J. Imaging Sci. & Tech.* **49**(1), 96-104 (2005).
- [Has09] Has M & Newman T. Color management: current practices and the adoption of a new standard. International Color Consortium. 17 Feb 2009. <http://www.color.org/wpaper1.xalter>.
- [Hunt57] Hunt R. The Reproduction of Colour (6th ed.) John Wiley & Sons Ltd. Chichester, England: (1957).
- [Hunt74] Hunt R, Pitt I & Winter L. The preferred reproduction of blue sky, green grass and Caucasian skin in colour photography. *J. Photographic Science*, **22**, 144-149 (1974).
- [Imai98] Imai F, Tsumura N, Haneishi H & Miyake Y. Improvements of incomplete chromatic adaptation model of facial pattern images. *J. Imaging Sci. Tech.* **42**(3): 243-268 (1998).
- [Jiang03] Jiang X. Estimation of illuminants from color signals of illuminated objects. PhD Thesis. Rochester Institute of Technology: Rochester, New York (2003).
- [Katz35] Katz D. *The World of Colour*. Trubner & Co., London, (1935).

- [Kuang05] Kuang J, Jiang X, Quan S & Chiu A. A psychophysical study on the influence of factors of color preference in photographic color reproductions. Proc. of SPIE-IS&T Electronic Imaging, 5668: 12-19 (2005).
- [Newhall57] Newhall S, Burnham R & Clark J. Comparison of successive with simultaneous color matching. J. Optical Society of America. 47(1), 43-56 (1957).
- [Phillips08] Phillips J, Burns P, Fredericks E & Rosen M, Image Quality and Cost Assessment of Print-on-Demand Books from Web-Based Vendors Offering One-Off Printing, Proc. of 24th International Conference on Digital Printing Technologies and Digital Fabrication 2008:450-454 (2008).
- [Pinney63] Pinney J and DeMarsh L. The study of colour reproduction by computation and experiment. J. Photographic Science, 11, 249-255 (1963).
- [Rosen06] Rosen MR, "Spectral Imaging" in Focal Encyclopedia of Photography, 4th Ed., Edited by M. Peres, Focal Press (2006).
- [Sharma03] Sharma G. Digital color imaging handbook. CRC Press. (2003).
- [Siple83] Siple P & Springer R. Memory and preference for colors of objects. Perception & Psychophysics, 34(4), 363-370 (1983).
- [Stelmack65] Stelmack R. The effect of hue and saturation on apparent distance judgements. MS Thesis. University of Windsor: Canada (1965).
- [Taplin04] Taplin L & Johnson G. When good hues go bad. IS&T CGIV Conference, Aachen Germany (2004).
- [Urban07] Urban P, Rosen MR & Berns R, Accelerating Spectral-Based Color Separation on the Neugebauer Hyperplane, J. Electronic Imaging, 16 (2007).
- [vonKries02] von Kries J. Chromatic adaptation. Festschrift der Albrecht-Ludwig-Universität. (Fribourg), (1902). [Translation: MacAdams D. Sources of Color Science, MIT Press, Cambridge, (1970)].
- [Weidner81] Weidner V & Hsia J. Reflection Properties of Pressed Polytetrafluoroethylene Powder, J. Opt. Soc. Am. 71(7): 856-861 (1981).
- [Yano97] Yano T & Hashimoto K. Preference for Japanese complexion color under illumination. Color Res. Appl. 22(4), 269-274 (1997).
- [You08] You J & Chien S. Saturation enhancement of blue sky for increasing preference of scenery images. IEEE Trans. on Consumer Electronics, 64(2): 762-768(2008).